A systematic review to assess the methodological quality of intervention research designed to increase physical activity in children with autism spectrum disorders

2013

Jolene Winn
University of Central Florida

A SYSTEMATIC REVIEW TO ASSESS THE METHODOLOGICAL QUALITY OF INTERVENTION RESEARCH DESIGNED TO INCREASE PHYSICAL ACTIVITY IN CHILDREN WITH AUTISM SPECTRUM DISORDERS

by

JOLENE B. WINN

A thesis submitted in partial fulfillment of the requirements
for the Honors in the Major Program in Health Sciences Pre-clinical
in the College of Health and Public affairs
and in the Burnett Honors College
at the University of Central Florida
Orlando, FL

Spring Term 2013

Thesis Chair: Dr. Jennifer Tucker

ABSTRACT

In comparison to normally developing children, many children with Autism Spectrum Disorders (ASD) do not possess the same opportunities to be physically active due to the impairments exhibited by their disorder. A systematic review using the Downs and Black checklist and the PEDro scale was conducted to assess the methodological quality of the literature on promoting physical activity in children with ASD. The following inclusion criteria had to be met: (1) subjects must include children with a clinical ASD diagnosis (2) the children have to be under the age of thirteen years old (3) the interventions must target physical activity; lastly, (4) they must be a relevant peer-reviewed English language study. The search was conducted using four electronic databases: MEDLINE, ERIC, PsycInfo, and CINHL with no restriction on the publication year. The following keywords were utilized: “Autism”, “ASD/ Autism Spectrum Disorder”, “Asperger”, “Pervasive Developmental Disorder” Those terms were paired with “physical activity”, “physical exercise”, “exercise”, “fitness”, “aerobic”, “swim”, “aquatic”, “jog”, “walk”, “recreational activity” Which were also paired with the terms “school age”, “child”, “toddler”, “preadolescent”. This multi-step search procedure occurred during February 2013.

The methodological quality of six studies was evaluated in February 2013.

Overall, the conclusive scores determined by the Downs and and Black checklist and the PEDro scale varied greatly. The scores reported by the Downs and Black checklist ranged from 19 to 21 on a 27-point scale. PEDro scale yielded scores ranging between two and six on a 10-point scale.

A vote count revealed that the exercise interventions increased the physical fitness, aquatic skills, social behaviors, and sensory integration children with ASD. In summary, the variation within the scores and the quality of the studies leads to a demand for future research. In order to adequately determine what exercise interventions effectively increase physical activity in children with ASD, future researchers should conduct randomized controlled trials in order to produce the highest quality of evidence.

DEDICATION

I would like to dedicate my thesis to my family. I would not be the woman that I am today without their continuous support and encouragement. I appreciate my parents for making huge sacrifices in their lives in order to provide my sister and me with a college education and comfortable lifestyle. Thank you to my Father, Jerome Winn, for always putting our family first and continuously encouraging me to pursue my dreams. Thank you for my mother, Margi Winn, for being an amazing role model. Continuing your education to increase your knowledge and professional skills is admirable. And lastly; thank you to my sister, Kalyn Winn. I appreciate all of the times you told me not stress out because you had faith that I will succeed. I don't know what I would do without any of you.

Additionally, I would like to thank my grandparents for continuously supporting me in everything that I pursue. Your inspirational conversations and past experiences serve as a true motivation for me.

I would like to give a special thank you to my mentor and thesis chair, Dr. Jennifer Tucker. I appreciate all of the time you spent assisting me with my proposal and thesis. You have truly provided me with knowledge that I will appreciate for a lifetime. This experience has not only contributed to the person that I am today, but it has encouraged me to continue my dreams of providing physical therapy to people diagnosed with autism spectrum disorders, a population that I have grown to love.

Furthermore, I would like to thank all of my committee members not only for their time and guidance, but also their constructive criticism. Your feedback has been crucial to my success, without you there would be no thesis. I appreciate all of the resources that you have provided me with in order to expand my research, skills, and horizons. Each have you have impacted and molded me during my undergraduate career which prepares me for my future endeavors. I am lucky to say that I have had the assistance of such knowledgeable and good hearted people.

TABLE OF CONTENTS

LIST OF TABLES

CHAPTER ONE Overview

The American Psychiatric Association (2000) defines Autism Spectrum Disorders (ASD) as an accumulation of intricate neurodevelopmental disabilities. In the United States, the most common pediatric diagnosis is ASD (Bhat, Landa, & Galloway, 2011). The prevalence of ASD is steadily increasing, as it is now estimated that every one out of 88 children are diagnosed (Centers for Disease Control and Prevention, 2012b). The resulting high incidence rates leads to many needs and health concerns for this population. Obesity is a main concern that affects America as a whole, including children with ASD (Centers for Disease Control and Prevention, 2012a). Several interventions have been explored to address this need, including: school programs, aerobic exercise, and hydrotherapy (Petrus et al., 2008). Furthermore, there is an increasing amount of evidence available for promoting physical activity in children with ASD. Nevertheless, the quality of evidence varies. The main purpose of this study is to assess the quality of the literature on promoting physical activity within this population of children. The overview of the literature will have clinical implications as the evidence may be incorporated into decision making on use of interventions to promote physical activity.

CHAPTER TWO Introduction

Autism Spectrum Disorders

Autism Spectrum Disorders (ASD) is defined as a group of complex neurodevelopmental disabilities that includes Autistic Disorder, Asperger's Disorder, Pervasive Developmental Disorder Not Otherwise Specified (PDD-NOS), Rett's Disorder, and Childhood Disintegrative Disorder. The main diagnostic characteristics commonly exhibited by these disorders are impairments in social interaction and communication skills, restricted interest, and may include redundant or stereotyped patterns of behavior and activities (American Psychiatric, 2000).

Overall, there are some important characteristics that distinguish each of these intricate neurodevelopmental disorders (See table 2.1).

Table 1: Characteristics of Autism Spectrum Disorders

	Autistic Disorder	Asperger's Disorder	Rett's disorder	PDD-NOS	CDD
Social and communication challenges	X	X	X	X	X
Delay or deviation of early language	X				
Poor gait and/or core movements			X		
Head growth deceleration			X		
Loss of prior developed hand skills			X		
Restricted, repetitive, and stereotyped interests	X	X			X
Distinctive pattern of developmental regression following at least 2 years of normal development					X
Only affects females			X		
Notable characteristics within the first year of life	X				
Does not meet the criteria for autism because of delayed age at onset, aberrant symptoms, sub threshold symptoms, or all of these				X	

PDD-NOS = pervasive developmental disorder not otherwise specified; CDD = childhood disintegrative disorder

Reference: American Psychiatric, A. (2000). Diagnostic and statistical manual of mental disorders (DSM-IV-TR).

Prevalence

The Centers for Disease Control (CDC, 2012) estimated that every 1 out of 88 children are diagnosed with ASD. It remains uncertain whether the increased amounts of children that are being diagnosed with ASD are reflected by a difference in diagnostics or inflation (American Psychiatric, 2000). A study conducted by Autism and Developmental Disabilities Monitoring (ADDM) Network, displays a greater prevalence in males (one in 54) than females (one in 252). ASD prevalence is highest in non-Hispanic white children (12.0 per 1,000), followed by non-Hispanic black children (10.2 per 1,000), and Hispanic children (7.9 per 1,000) as revealed by a combination of the data from the 14 sites. Over time, the prevalence of ASD has increased the most among Hispanic children and non-Hispanic black children (Autism and Developmental Disabilities Monitoring Network Surveillance Year 2008 Principal Investigators & Centers for Disease Control and Prevention, 2012)

Physical Activity

Since a vital component of a healthy lifestyle is physical activity, it is recommended that all children engage in physical activity, as well as get annual check-ups and screenings. In addition to promoting good physical health, regular physical activity also increases lean body mass, muscle, and bone strength. Furthermore, it fosters psychological well-being by acting as a stress reliever and increasing self-esteem and the capacity for learning. The guidelines for physical activity that apply to all children are moderate to vigorous activity for at least 60 minutes per day, or at least most days of the week (American Academy of Pediatrics,). By not fulfilling this requirement, many children increase their risk of becoming obese. In order to

prevent the incidence of pediatric obesity, the American Academy of Pediatrics (AAP) recommends that BMI measurements be completed once a year on all children and adolescents. A BMI is used to classify children as: Overweight (At or above 95th percentile of BMI-for-age), at risk for overweight (between 85th and 95th percentiles), normal weight (between 5th and 85th percentiles), and underweight (below 5th percentile of BMI-for-age). The AAP also recommends the promotion and support of a healthy lifestyle, such as healthy eating patterns and increasing physical activity. Screen time is also recommended as a daily maximum of two hours (Krebs & Jacobson, 2003)

Even though the AAP has straightforward recommendations and guidelines, majority of children are not engaging in enough physical activity. Trost et al. (2012) studied the physical activity patterns of 470 inner-city elementary school children. The physical activity recommendation of at least 60 minutes of moderate to vigorous exercise was met by less than one in four inner city school children. Otherwise stated, the current standards for physical activity were only met by 24.3%. The study indicated that low-income, predominantly African American and Hispanic children had much lower levels of physical activity than the national average (Trost et al., 2012). Nyberg, Nordenfelt, Ekelund, & Marcus (2009) administered a large study that evaluated the physical activity patterns of 1538 children 6- to -10 years of age. There was already a decline in physical activity levels at the age of 6 years, as implied by the results of the study. In addition, there were drastically low levels of physical activity of children during the evening and on weekends in comparison to weekdays. Overall, the physical activity level was found to be greater in boys rather than girls. The decrease in levels of physical activity among children is primary concern that remains an issue throughout the United States.

Physical Activity in Children with ASD

Unlike typically developing children, numerous children with ASD do not have the opportunity to participate in physical activity because of their disability; therefore, they are not given the same choices to be active. For example, the opportunity to become involved in a program that provides physical activity may be restricted because of one's behavioral issues. The Special Olympics and other segregated programs are available; however, the child's decision to participate is determined by social, cognitive, and cultural factors. As a result, the participation may be limited. The engagement in participation of physical activities is a concern for children with ASD because of the social deficits investigated (C. Y. Pan & Frey, 2006).

A study conducted by Pan and Frey (2006) tested the levels of physical activity of nine middle schools. Physical education lessons were provided by the schools two times per week for 45 minutes. They were able to assess the physical activity levels and patterns of the participants and the outcome of the study revealed that children with ASD seemed to be less physically active than children without disabilities. In addition, the findings revealed that children with ASD had significantly lower steps/min than their peers without disabilities. According to Pan, Tsai, and Hsieh (2011) students in middle school who were diagnosed with ASD were less active in physical activities than those in lower grade levels. This was also supported by MacDonald, Esposito, & Ulrich (2011) who assessed the physical activity patterns of 72 children between the ages of 9-18 that were diagnosed with autism. The study revealed that physical activity patterns decline overtime as children with autism age. One of the main concerns about this assessment was that overall time spent doing vigorous physical activity was only one minute. Furthermore,

the study revealed a decrease in physical activity after school, which suggests a need for afterschool programs.

Consequently, research suggests that the lack of physical activity exhibited by all children is contributing to the current obesity issues in America. Furthermore, the lack of physical activity among the ASD population can be said to be contributing to the overall increasing obesity rates of all children. Due to the high incidence, increasing obesity rates, and lack of physical activity among children with ASD, high quality research is warranted in order to determine which intervention most effectively increases physical activity within this population. By assessing the methodological quality of current research, the strengths and limitations of a study will be addressed and interpreted in order to reveal the true effectiveness of an intervention. Any methodological deficiencies of a study may alter the results. Furthermore, it will allow researchers, healthcare practitioners, and readers to become aware of potential bias within the results of studies. In conclusion, this approach will adequately determine which study design produces the most valid and reliable information in order to establish which intervention is most effective.

CHAPTER THREE Methodology

Study design

A systematic review using the Downs and Black checklist and the PEDro scale was used to assess the methodological quality of interventions used to increase physical activity in children with ASD. The findings will have clinical implications that may be integrated into the decision making for future interventions used to promote physical activity.

Instruments

In 1998, Downs and Black developed the 27 item Methodological Quality Checklist. This systematic review will rate the studies using an amended Downs and Black checklist. Twenty-six 'yes'-or-'no' questions will be scored totaling up to 26 possible points. In this review, the questions were categorized under 4 sections: Reporting (10 items), External validity (three items), Study bias (seven items), and Confounding and selection bias (six items) (Downs & Black, 1998). The Downs and Black checklist is a significant tool that will assess the reliability and validity of past studies that have explored interventions used to increase physical activity in children with ASD.

The PEDro scale, another methodological assessor, will also be used to evaluate the studies. The PEDro scale, an adopted rating measurement used by the Physiotherapy Evidence Database, consists of 11 questions that formulate a total of 10 points. Based on the criteria, the rating will provide an evaluation of the methodological rigor associated with a trial. The validity

is presumed to be greater with higher scoring trials (Centre for Evidence-Based Physiotherapy at The George Institute for Global Health, 1999).

Screening Procedure/ Inclusionary Criteria

Furthermore, a vote count was also completed in order to determine the overall effect that exercise interventions have on the outcomes measured within the inclusive studies. To be considered significant, the study must have reported a key outcome with a p-value equal to 0.05 or less. Several of the outcomes measured within the six studies were grouped into the following categories: physical fitness, aquatic skills, social competence, antisocial behavior, and sensory integration/stereotypic behaviors.

In preparation of the systematic review, several training sessions were held by three faculty members at the University of Central Florida. For the purpose of ensuring accurate interpretation, 75% of the articles were scores by myself and one other faculty member. The scores were compared. Any score discrepancies were discussed and both raters came to an agreement about the score. The purposes of the training sessions were to increase the internal consistency of the scores and to improve the quality of the systematic review.

Search procedures

The search was conducted using four electronic databases: MEDLINE, ERIC, PsycInfo, and CINHL with no restriction on the publication year. The following keywords were utilized: “Autism”, “ASD/ Autism Spectrum Disorder”, “Asperger”, “Pervasive Developmental Disorder” Those terms were paired with “physical activity”, “physical exercise”, “exercise”, “fitness”,

"aerobic", "swim", "aquatic", "jog", "walk", "recreational activity" Which were also paired with the terms "school age", "child", "toddler", "preadolescent". This multi-step search procedure occurred during February 2013.

Inclusion Criteria

To be included in this systematic review, the following four inclusion criteria had to be met: (1) Subjects must include children with a clinical ASD diagnosis (2) the children have to be under the age of thirteen years old (3) the interventions must target physical activity; lastly, (4) they must be a relevant peer-reviewed English language study. Any articles not meeting all four criteria were excluded.

Participants

Conclusively, the six studies recruited 210 people and provided intervention to a total of 176 participants. Two studies that reported participant drop outs account for the loss of 20 individuals. The gender of the 11 participants that dropped out of the intervention was not reported by Wuang, Wang, Huang & Su (2010). As a result, there were 130 males and 35 females that were reported collectively. In other words, the male to female ratio was 26 to seven. The gender distribution is true to the male to female ratio reported for the ASD population since autism is considered to be more prevalent in males (Autism and Developmental Disabilities Monitoring Network Surveillance Year 2008 Principal Investigators & Centers for Disease Control and Prevention, 2012).

Overall, the combination of the studies totaled 131 children with Autism, 15 with Asperger's disorder, three with Pervasive Developmental Disorder Not Otherwise Specified, two with Cerebral Palsy (CP), three with Developmental Delay (DD), two with Down Syndrome (DS), two with Myelomeningocele (MCC), one with Nonverbal learning disorder, one with Otopalataldigital syndrome, one with intellectual disability, and 15 non-ASD. Generally, this does reflect the target population, children with ASD, since the majority of the participants evaluated within the studies was clinically diagnosed with an ASD.

Exercise interventions

The six studies under review reported the following primary exercises: water aerobics (n=3), running/jogging (n=1), and horseback riding (n=2). Although all of the reviewed studies incorporated physical activity within their experiment, the intentions of their measurement were diverse. For instance, they may have used physical activity for secondary benefits, depending on the researcher's motive. Some of the main outcomes that the physical activity measured were cardiorespiratory endurance, physical condition, academic engagement, attention, as well as social and communication skills.

CHAPTER FOUR Results

In summary, the conclusive scores determined by the Downs and and Black checklist and the PEDro scale were diverse. The scores reported by the Downs and Black checklist ranged from 19 to 21 on a 27-point scale. See Table 3 for details. As displayed in Table 4, the PEDro scale yielded scores ranging between two and six on a 10-point scale. There were four significant outcomes and one non-significant outcome as supported by the vote count in Tables 5-9.

Table 2: Experimental Overviews

Study	Participant Characteristics	Experimental Design	Targeted Measures	Intervention Frequency	Outcomes
Fragala-Pinkham, Haley, and O'Neil (2008)	20 recruited; four dropped out; 16 children aged 6 to 12 years old completed the study; two with HFASD, one with Autism and ADD, three with PDD-NOS, two with CP, two with DD, two with DS, two with MCC, one with Nonverbal learning disorder, one with Otopalataldigital syndrome with ADD	Non-randomized A-B group design; one group	Cardiorespiratory endurance, muscle strength, motor skills, and heart rate	14 weeks; Two 50min aquatic sessions per week with at least two days between sessions	Increased cardiopulmonary endurance, most children improved ability to exercise longer in their target HR zones; No significant improvement in motor skills or muscle strength of abdominal and lower extremities
C. Y. Pan (2010)	16 boys aged 6 to 9 years old; eight with HFASD and eight with Asperger disorder	Within-participant repeated-measures design	Aquatic skills and social behaviors	21 weeks (10 weeks WESP, 10 weeks control, one week transition); The 10 week WESP was 20 sessions; two 90min sessions per week	The WESP improved the aquatic skills in four out of five stages and decreased antisocial behavior problems; no effect on reducing stereotypic behaviors

Bass, Duchowny, and Llabre (2009)	34 participants; experimental group (n=19) : two girls and 17 boys between 5-10 years old; one with Asperger disorder, six with mild autism, 10 with moderate autism, two with severe autism; Control group (n=15): three girls and 12 boys between 4-10years old; one with Asperger disorder, five with mild autism, six with moderate autism, three with severe autism; six participants dropped out of the experimental group and three from the control group	Randomized Controlled Trial (RCT)	Social functioning	12 weeks; one hour weekly therapeutic horseback riding sessions	The experimental group improved in sensory integration and directed attention, improved social motivation and sensory sensitivity, and decreased inattention and distractibility; no significant effect in fine motor/ perceptual, social cognition, and social awareness
Oriel et al. (2011)	Nine participants; seven males and two females between 3 and 6 years old; seven with autism, one with intellectual disability, and one with DD	Randomized, within-subjects crossover design	The number of correct and incorrect academic responses, stereotypic behaviors, and on-task behavior	Six weeks; three weeks of treatment condition and three weeks of control condition	Statistically significant Improvements in correct responding following 15 min run; No statistical improvements observed in on task time and reduction of stereotypic behaviors

C. Pan (2011)	30 Participants; Experimental group (n=14); seven with ASD (all male, five with autism and two with Asperger disorder; seven non-ASD (six females, one male); Control group (n=16); eight with ASD (all male, five with autism and three with Asperger disorder; eight non-ASD (four females and four males)	Controlled, single blinded, within-participant repeated-measures design	Aquatic skills and physical fitness of children with ASD and their siblings	32 week program, 14 weeks aquatic program, 14 weeks control, four weeks assessment and transition; Experimental group (first 14 weeks of aquatic program), Control group (second phase of another 14 weeks of aquatic program); 28 sessions (2—60min sessions/week)	Significant increase in all subjects of physical fitness and aquatic skills except subtest of body composition; Significant increase in muscular strength/endurance and more advanced water skills in experimental group compared to control group;
Wuang et al. (2010)	71 participants with ASD; split into two groups: group A (35), group B (36); 11 children (5 from A 6 from B) dropped out; 60 completed study; 13 girls and 47 boys	Controlled, single blinded, within-participant repeated-measures design	Gross and fine motor skills, sensory integration	44 week SDHRP; 20 weeks SDHRP, 20 weeks control, 4 weeks assessment and transition; One hour sessions 2 times per week (40 sessions per group)	Significant increases in gross motor subtests in group A; All TSIF scores increased; No significant effects seen in fine motor function

HFASD = high-functioning autism spectrum disorder; PDD-NOS = pervasive developmental disorder not otherwise specified; CP = cerebral palsy; DD = developmental delay; DS = down syndrome; MCC = myelomeningocele; ADD = attention deficit disorder, WESP = water exercise swimming program, SDHRP = simulated developmental horse-riding program, TSIF = test of sensory integration function

Table 3: Results of the Downs and Black checklist evaluations

	Fragala-Pinkham, Haley, & O'Neil (2008)	C. Y. Pan (2010)	Bass, Duchowny, and Llabre (2009)	Oriel et al. (2011)	C. Pan (2011)	Wuang et al. (2010)
Reporting	10/11	10/11	9/11	9/11	10/11	9/11
External validity	3/3	1/3	3/3	1/3	3/3	3/3
Internal validity - bias	6/7	6/7	5/7	5/7	6/7	6/7
Confounding and selection bias	1/6	3/6	4/6	5/6	2/6	1/6
Total	**20/27**	**20/27**	**21/27**	**20/27**	**21/27**	**19/27**

Table 4: Results of the PEDro scale evaluations

	Fragala-Pinkham, Haley, & O'Neil (2008)	**C.Y. Pan (2010)**	**Bass, Duchowny, and Llabre (2009)**	**Oriel et al. (2011)**	**C. Pan (2011)**	**Wuang et al. (2010)**
Eligibility criteria specified (not scored)	Yes	Yes	Yes	Yes	Yes	Yes
Subjects were randomly allocated to groups	No = 0	No = 0	Yes = 1	Yes = 1	No = 0	No = 0
Allocation was concealed	No = 0	No = 0	No = 0	No = 0	No = 0	No = 0
The groups were similar at baseline regarding the most important prognostic indicators	No = 0	Yes = 1	Yes = 1	No = 0	Yes = 1	Yes = 1
There was blinding of all subjects	No = 0	No = 0	No = 0	No = 0	Yes = 1	No = 0
There was blinding of all therapists who administered the therapy	No = 0	No = 0	No = 0	No = 0	No = 0	No = 0
There was blinding of all assessors who measured at least one key outcome	Yes = 1	Yes = 1	No = 0	Yes = 1	No = 0	Yes = 1
Measures of at least one key outcome were obtained from more than 85% of the subjects initially allocated to groups	No = 0	Yes = 1	Yes = 1	No = 0	Yes = 1	Yes = 1
All subjects for whom outcome measures were available received the treatment or control condition as allocated or, where this was not the case, data for at least one key outcome was analyzed by “intention to treat”	No = 0	Yes = 1	Yes = 1	Yes = 1	Yes = 1	No = 0
The results of between-group statistical comparisons are reported for at least one key outcome	No = 0	Yes = 1	Yes = 1	Yes = 1	Yes = 1	Yes = 1
The study provides both point measures and measures of variability for at least one key outcome	Yes = 1	Yes = 1	Yes = 1	No = 0	Yes = 1	Yes = 1
TOTAL	**2/10**	**6/10**	**6/10**	**4/10**	**6/10**	**5/10**

Table 5: Vote count for physical fitness

	Significant	Not significant
Fragala-Pinkham, Haley, & O'Neil (2008)	X	
Bass, Duchowny, and Llabre (2009)		X
C. Pan (2011)	X	
Wuang et al. (2010)	X	
SCORE	**3**	**0**

Table 6: Vote count for aquatic skills

	Significant	Not significant
C. Y. Pan (2010)	X	
C. Pan (2011)	X	
SCORE	**2**	**0**

Table 7: Vote count for anti-social behavior

	Significant	Not significant
C. Y. Pan (2010)	X	
Bass, Duchowny, and Llabre (2009)	X	
SCORE	**2**	**0**

Table 8: Vote count for social competence

	Significant	Not significant
C. Y. Pan (2010)		X
Bass, Duchowny, and Llabre (2009)		X
SCORE	**0**	**2**

Table 9: Vote count for sensory integration/ stereotypic behaviors

	Significant	**Not significant**
Bass, Duchowny, and Llabre (2009)	X	
Oriel et al. (2011)		X
Wuang et al. (2010)	X	
SCORE	**2**	**1**

CHAPTER FIVE Discussion

Downs and Black checklist results

Overall, the evaluated studies received high scores on the first subscale, Reporting. The high scores were reflected by the detailed reports given by the inclusive studies. The Majority of the studies were clear about their main objective, outcomes to be measured, interventions of interest, main outcomes, estimates of variability, loss of participants, and probability findings. However, all of the reviewed studies did lose one point for not reporting adverse effects that may have been a consequence of the intervention. If any adverse effects occurred during the study then it could have altered the results. For instance, if a participant became ill secondary to the intervention and it was not reported, future researchers or healthcare practitioners will not be aware of potential adverse effects. On the other hand, various studies neglected to provide a list the principal confounders. Instead of just stating that the participants were similar at baseline, all studies should have also stated the specific characteristics, such as age, gender, and severity, for each and every participant. Failure to account for a confounding factor restricts the reader's ability to have any alternate interpretations, which in turn increases bias within the study.

The consecutive subscale on the Downs and Black checklist is external validity, which attempts to address whether the outcomes of the study were representative and generalized to the population from which the participants were derived. The criteria for external validity were met by all of the evaluated studies except for C. Y. Pan (2010) and Oriel, George, Peckus, and Semon (2011) who only received one out of three possible points. The lack of clarity about which schools the participants came from in both studies was reflected in a lower score. The attendance

of a particular type of school may give some children an advantage over the others since there is not a national standardized physical education program for all schools.

The following subscale, Internal validity – bias, addresses whether blinding was attempted, data dredging was stated, appropriate statistical tests were used, compliance within interventions, and the accuracy of the outcome measures. This subscale resulted in only a one point difference between numerous studies. Bass, Duchowny, and Llabre (2009) was the only study that did not make an attempt to blind either the assessors or participants during the intervention. Consequently, there could have been possible bias not accounted for within the results of the study. There was a lack of certainty about who was blinded in the study conducted by C. Pan (2011). The study claims to be a single blinded study, yet it failed to mention blinding of assessors or therapists administering the intervention. There is a discrepancy about whether or not the participants were aware that they were involved in an intervention. However, a point was awarded for blinding of the subjects since it was stated in the title that it was a single-blinded study.

The last subsection, Confounding and selection bias, yielded the largest variety of scores. The differences among the scores were mainly due to the contrasting experimental designs. Higher quality experiments have less selection bias because they use a randomization assignment of participants as a part of their methodology. Unfortunately, only two out of the six studies used a randomization method to assign participants to groups. Although the two studies randomly assigned their participants to either the experimental group or the control group, they failed to state who performed the randomization or how it took place. If the researcher who determined

the eligibility of a subject also organized the randomization process, there could have been unconscious selection bias when assigning the participants into groups. The use of a randomizing computer program or blinding of the person assigning the groups would greatly reduce any possible selection bias that can occur during the randomization process. Nevertheless, there was some apparent selection bias within the studies conducted by C. Y. Pan (2010), C. Pan (2011) and Oriel et al. (2011) because the recruited participants were not representative of the population that was targeted for their study. For instance, C. Y. Pan (2010) attempted to measure the effects of a water exercise swimming program on aquatic skills and social behaviors in children with ASDs. Yet, the participants only met the inclusion criteria if they were diagnosed with Asperger disorder or mild and/or high functioning autism. Since ASD consist of a broad spectrum it is unreasonable to claim that solely high functioning autism is reflective of the entire population of children with ASD. This was also evident in the study conducted by Oriel et al. (2011), which measured the effects of aerobic exercise on academic engagement in young children with ASD. Although the participants of the study meet the educational criteria for ASD, only seven out of nine children were clinically diagnosed with formal autism. Researchers should take into consideration all aspects of the condition instead of generalizing. As a result, there will be less misconceptions and more precise findings.

PEDro scale results

Since the PEDro scale is primarily used to assess a Randomized Controlled Trial (RCT), many of the evaluated studies resulted in low scores due to their experimental design. As expected, the highest score obtained by Bass, Duchowny, and Llabre (2009) was the only RCT evaluated. Bass, Duchowny, and Llabre (2009) and Oriel et al. (2011) were the only evaluated studies that randomly allocated their subjects into groups. However, as stated earlier, they failed to conceal the allocation which may have led to the influence of a particular group. For example, the sequencing of treatments to be provided could have been changed which may have affected the results. This is problematic because it could generate systematic biases in an otherwise random distribution.

Furthermore, bias within treatment outcomes is a concern for the studies conducted by Oriel et al. (2011) and Fragala-Pinkham, Haley, and O'Neil (2008) because the participants were not similar at baseline. For instance, the 16 children involved in the intervention executed by Fragala-Pinkham, Haley, and O'Neil (2008) were all of different diagnoses that may have had different mobile abilities. Similarly, the Oriel et al. (2011) study had two children primarily diagnosed with intellectual disability and developmental delay in comparison to the other seven children that had a primary ASD diagnosis. The severity of each child's condition could have affected their ability to fully partake in all aspects of the intervention. Group outcomes could be expected to differ because of the dissimilarity in key prognostic variables at baseline.

Although five out of six studies had either blinding of subjects or blinding of the assessors who measured the key outcomes, it is worrisome that none of the inclusive studies

attempted to blind the therapists who administered the therapy. The blinding of the therapists ensures that the effects of the interventions are not due to the therapists' excitement or lack of excitement for either the treatment or control group.

In comparison to the other studies, Oriel et al. (2011) and Fragala-Pinkham, Haley, and O'Neil (2008) were the only studies whose interventions did not yield key outcome measures collected by more than 85% of their participants. The study conducted by Oriel et al. (2011) had such a small number of participants that overall their key outcomes measured did not meet the criteria on the PEDro scale. Alternatively, the loss of participants in the Fragala-Pinkham, Haley, and O'Neil (2008) study resulted in only 80% completion of the full intervention, which in turn negatively impacted the key outcome. Moreover, the loss of patients that was not included within the analyses of the studies conducted by Wuang et al. (2010) and Fragala-Pinkham, Haley, and O'Neil (2008) justifies that the main conclusions of the study were based on treatment rather than intention to treat. As a result, it can be inferred that large sample sizes and the reporting of all data within studies are crucial to producing a high quality study.

With regards to between-group statistical comparisons, all of the inclusive studies reported at least one key outcome, except for Fragala-Pinkham, Haley, and O'Neil (2008). Since the study only had one group, they were unable to report any key outcomes of between-group statistical comparisons. For this reason, the overall score of the study suffered. Without a control group, there is no foundation for knowing whether a specific result is due to the variable being tested or other factors.

Experimental designs

A strong experimental design is the most crucial factor impacting the methodological quality of a study. It is evident that the randomized and blinded studies yielded greater results because there was less opportunity for bias to be produced within the studies. Bass, Duchowny, and Llabre (2009) and C. Pan (2011) attained the highest scores on both the Downs and Black checklist and the PEDro scale. The experimental design used by Bass, Duchowny, and Llabre (2009) was a RCT. Although the study received one of the highest scores, it was the only study that failed to attempt any blinding. Blinding of the subjects, assessors, or therapists may substantially decrease internal validity bias. Oriel et al. (2011) conducted a within-subjects crossover design that also used a randomization process. However, the subjects were not individually assigned to groups; instead two of four early intervention classes were randomly assigned to either a treatment condition or a control condition. The randomization process implemented by both researchers minimized selection and allocation bias in the treatment assignment. Furthermore, the crossover design allowed both groups to take part in both conditions for the same amount of time. For instance, after randomly being assigned to either the treatment or control group, the participants took part in the intervention under their assigned condition for three weeks then received the opposite condition for the subsequent three weeks of the study. This design created balance between the two groups since both groups receive the same amounts of treatment over the same period of time. In addition, it allowed the participants to serve at his or her own control. C. Pan (2011) and Wuang et.al (2010) both conducted controlled, single blinded, within-participant repeated-measures design. Since the experiments were controlled, there was a control group which allowed comparisons to be made when the

results of the study were analyzed. C. Pan (2011) blinded the participants, which means that the participants did not know whether or not they were in a treatment or control group. As a result, the research suggests that the effects of the treatment were not altered by the placebo effect. The study conducted by Wuang et al. (2010) blinded the assessors, which means that the people conducting the assessments were blind to the child's group status. In particular, there were two pediatric occupational therapists who administered the assessments to the children during three different evaluations. Fragala-Pinkham, Haley, & O'Neil (2008) also blinded the assessors that administered the outcomes of the results. In addition, C.Y. Pan (2010) also blinded the assessors for one of the main key outcomes. The blinding of the assessors makes the methodological design of the study higher quality because there is no bias within the assessments which increases the validity of the results. Furthermore, the within-participant repeated-measures design, used by C. Pan (2011), Pan (2010), and Wuang et al. (2010), collected the same number of measurements from every subject in both the treatment and control group. Consequently, the researchers were able to study changes made within each participant over a long period of time. The consistency of a measure and small number of groups allow experiments to run in an orderly manner. Disadvantages of the within-participant repeated-measures design is that participants may not be able to make it to every session of the experiment or even complete the entire intervention at all, which could negatively affect the outcome of the study. For example, Wuang et al. (2010) had 11 out of 71 children that did not complete the entire intervention. As a result they were only able to analyze a full data set of 60 people. However, the loss of participants did not have a significant effect on the intervention outcome due to their large sample size. Overall, the study administered by Fragala-Pinkham, Haley, & O'Neil (2008) had the poorest

experimental design. The implemented A- B group design only consisted of one group. As a result, they were unable to do any between group statistical analyses. Moreover, the study lacked a sufficient number of participants to produce any valuable and trustworthy data.

In summary, the contrasting scores resulting from both the Downs and Black checklist and the Pedro scale were due to a difference in methodology. The inclusive studies received overall higher scores on the Downs and Black checklist. In comparison to the PEDro scale, the Downs and Black checklist provides an increased amount of questions that inquire more detail. This could be a possible explanation for the difference in scores since the Downs and Black checklist presents more opportunity to accumulate points.

Experimental significance

In order to determine the effectiveness of interventions used to increase physical activity in children with ASD, a vote count was conducted to determine whether or not the overall outcomes of the exercise interventions were deemed significant.

Table 10: Summary of the (a) measurements (b) significant outcomes

Study	Measurement	Significant Outcome	P-value
Bass, Duchowny, and Llabre (2009)	Social Responsiveness Scale and the Sensory Profile questionnaire	Increased social functioning and sensory integration	$P < 0.01$
C. Y. Pan (2010)	Aquatic skills: Humphries' Assessment of Aquatic Readiness; Social behaviors: School Social Behavior Scales	Improvement in aquatic skills; decreased antisocial behavior	$P < 0.01$
C. Pan (2011)	Aquatic skills: Humphries' Assessment of Aquatic Readiness; Physical fitness: Progressive Aerobic Cardiovascular Endurance Run (PACER), Curl-up test, Sit-and-reach test	Increase in aquatic skills and all of the physical fitness subtests except for body composition	$P < 0.05$
Oriel et al. (2011)	Observation	Increased academic performance	$P < 0.05$
Wuang et al. (2010)	Motor function: Bruininks-Oseretsky Test of Motor Proficiency	Improvements in gross motor subtests	$P < 0.01$
Fragala-Pinkham, Haley, and O'Neil (2008)	Half-mile walk/run	Increased cardiopulmonary endurance	$P < 0.001$

Overall, the vote count suggests that exercise interventions are an effective way to increase the physical fitness of children with ASD. This was supported by Fragala-Pinkham, Haley, and O'Neil (2008), Pan (2010), and Wuang et al. (2010). Fragala-Pinkham, Haley, and O'Neil (2008) conducted their intervention twice per week for 14 weeks with at least two days between sessions. The program consisted of a 3-5min pool warm up session, 20-30min of aerobic exercise, 5-10min of strength training, and 3-5min of cool down and stretching. Participants were measured twice at baseline before the intervention, and once at the end of the 14 week intervention. A statistically significant increase in cardiopulmonary endurance was revealed by their study. The participants displayed apparent reductions in the time to complete a half-mile walk/run. In contrast, C. Pan (2011) conducted a 32 week program, 14 weeks aquatic program, 14 weeks control, and four weeks assessment and transition. Participants were assessed three times: once at baseline, a second time after the first 14 weeks of aquatic program or regular treatment/activities, and third time after another 14 weeks. The program constituted of 28 sessions (two—60min sessions/week) including: 10min of structured social & floor warm up activities, 35min practice of individual or partnered treatment goals, 15min of group games/activities, and the last 10min of cool down activities. As stated in Table 10, the significant outcomes of the physical fitness of the participants were measured using a Progressive Aerobic Cardiovascular Endurance Run (PACER), Curl-up test, and Sit-and-reach test. The PACER consisted of a 16-m multi-stag shuttle run. Participants were prompted to run the distance for as long as possible at a specified pace which steadily increased each minute. The curl up test was scored by the number of curl ups a participant accomplished, or when 75 curl ups were completed. Generally, the study revealed that the participants in the experimental group had a

significant increase in aquatic skills, and all of the physical fitness subtests except for body composition. The treatment gains of aquatic program were maintained for 14 weeks on the function of muscular strength/endurance in children with ASD. The experimental group also showed a great increase in muscular strength/ endurance and more advanced aquatic skills than the controlled group. Lastly, the study conducted by Wuang et al. (2010) consisted of a 44 week Simulated Developmental Horse-Riding Program (SDHRP). There was 20 weeks of SDHRP, 20 weeks control, and four weeks of assessment and transition. Participants were assessed three times, once during the first week, a second time during the 22nd and 23rd week, and then in the final 44th week. Each treatment was composed of three sessions, each preceded with a warm-up. The first session involved simple limb movements and mat exercises. The second session focused on mounted exercises: ride the simulated horse in different positions (sitting, prone, lying). The third session consisted of playing a game on the simulated horse. Conclusively, the study revealed that group A had significant improvements in gross motor subtests, which included: running speed/agility, bilateral coordination, strength, and balance. Furthermore, the gross motor subtests in group A implied that the treatment gain of SDHRP could be maintained for at least 23-24 weeks.

Similarly, the score displayed in the vote count for aquatic skills suggests that aquatic exercise interventions are an effective way to increase aquatic skills of children with ASD. This was supported by both of Pan's studies conducted in 2010 and 2011. C. Y. Pan (2010) administered a 21 week intervention (10 weeks of the Water Exercise Swimming Program (WESP), 10 weeks control, and one week transition). The 10 week WESP was 20 sessions (two 90min sessions per week) and consisted of (1) floor activities (2) one-to-two instruction (3)

group activities (4) cool down activities. Each participant was measured three times: once at entry which served as a baseline, a second time after 10 weeks of WESP or regular treatment/activity, and a third time after another 10 weeks. Identical to the measure used to in Pan's other study conducted in 2011, the Humphries Assessment of Aquatic Readiness checklist (HAAR) was used to evaluate the aquatic skills of each participant. The HAAR consists of five stages: Mental adjustment (five items), Introduction to water environment (10 items), Rotations (three items), Balance and control (eight items), and lastly, Independent movement in water (six items). C. Y. Pan (2010) revealed an improvement in the subject's aquatic skills in four out of the five stages of the HAAR checklist. In addition, the effects of the WESP could be maintained for at least 10 weeks. The other study administered by C. Pan (2011) revealed that the children in the experimental group had significant increases in aquatic skills in comparison to the control group. Moreover, the gains of the treatment in children with ASD were solely maintained for 14 weeks on function of muscular strength/endurance.

Additionally, the vote count revealed that exercise interventions effectively decreased anti-social behavior children with ASD. This outcome was evident in the study conducted by Bass, Duchowny, and Llabre (2009) which consisted of one hour therapeutic horseback riding sessions per week over the span of 12 weeks. Each session was comprised of: mounting/dismounting of the horse, warm-up stretching exercises, riding skills, mounted games, and horsemanship activities. Participants were measured once before the 12 week intervention and a second time after completion of the intervention. The Social Responsiveness Scale (SRS) and the Sensory Profile questionnaire were used to measure the social behavior of the participants. The SRS was based off of five subscales including social awareness, social

communication, social motivation, autistic mannerisms, and social cognition. The Sensory Profile questionnaire was also focused on five subscales: fine motor/perception, sensory seeking, attention and distractibility, sensory sensitivity, and sedentary. Overall, the results of the experimental group revealed an improvement in social motivation, directed attention and overall social functioning. C. Y. Pan (2010) also measured anti-social behavior in children with ASD by using School Social Behavior Scales (SSBS) which assesses anti-social behavior, as well as social competence. Anti-social behavior was defined as being antisocial/aggressive, defiant/disruptive, and hostile/irritable. An assessment was conducted by the classroom teachers who were blind to the treatment condition of the participants. The typical anti-social behavior problems present in most children with ASD decreased as a result of the WESP.

In contrast to the other vote counts, there was no significant outcome for increasing social competence in children with ASD. C. Y. Pan (2010) noted that social competence is comprised of self/management/compliance, peer relations, and academic behavior. As reported by Pan (2010), the WESP did not increase social competence behaviors. Bass, Duchowny, and Llabre (2009) did not report significant effects in social cognition or awareness, two important components of social competence.

Furthermore, the vote count for sensory integration/ stereotypic behaviors revealed that exercise interventions are an effective way to increase sensory integration in children with ASD. Oriel et al. (2011) reported no significant statistical outcomes for reduced stereotypic behaviors. However, Bass, Duchowny, and Llabre (2009) and Wuang et al. (2010) both reported an improvement in sensory integration as a result of physical activity.

An increase in academic performance was also a significant key outcome measure as revealed by Oriel et al. (2011). However, since it was the only intervention that measured academic performance, a vote count could not be performed for it since there were no other studies to compare it to.

Limitations

The original intent of the systematic review was to only assess RCTs. However, due to the lack of this type of experimental design, all types of designs were evaluated. A limitation of this study was the lack of clarity within the reports. Without directly stating specifics, the researcher makes it difficult for the reader to decipher what exactly occurred or what the true intentions of the study were. Inherently, these challenges decrease the quality of the study.

CHAPTER SIX Conclusion

In conclusion, future research is warranted in order to determine what exercise interventions effectively increase physical activity in children with ASD. Although a number of the inclusive studies scored moderately on the Downs and Black checklist, the vast majority had substantial deficiencies when scored on the PEDro scale. The heterogeneity among the methodological quality of the inclusive studies reflected in their scores. For instance, the RCT earned top scores in both assessments in comparison to the A-B group design which received some of the lowest scores reported. It is recommended that future researchers review the criteria of a quality experimental design before implementing an intervention. In return, a larger body of research will exist as RCTs, the golden standard. Majority of the studies failed to randomize the participants, one of the main attributions of an RCT. Consequently, it is necessary for researchers to improve their methodological designs in future studies. All studies should be of high quality in order to convey unbiased information to decision-makers, future researchers, and practitioners. The evaluated studies frequently lacked clarity when reporting or describing various parts of their study which may have increased bias and restricted the reliability and validity of their conclusions. In particular, all of the studies failed to report adverse effects. Future studies need to report whether any adverse effects occurred in order to improve the quality of the study and to inform others about any possible safety precautions that may need to be implemented when re-conducting the study. Furthermore, all studies need to have a control group in order to provide more valid results reduce possible error. Another important issue that should be addressed in every study is a large number of participants. As exhibited by Oriel et al. (2011), a small number

of participants make it difficult to obtain significant results and furthermore these results may not be reflective of a larger population.

According to the vote count, the outcomes of the exercise interventions revealed that physical activity increases the physical fitness, aquatic skills, social behaviors, and sensory integration in children with ASD. An overview of the vote count suggests that increasing physical activity in children with ASD is likely beneficial for their overall health. These clinical applications provide a basis for future researches, practitioners, and even families with children that have ASD. The evidence may be incorporated into decision making on use of interventions to promote physical activity in children with ASD.

Overall future researchers need to give careful attention to the design and quality of their study in order to produce unbiased, effective results.

REFERENCES

American Academy of Pediatrics.Promoting physical activity in 3rd edition guidelines, pocket guide, tool & resource kit. Retrieved November 21, 2012, from http://brightfutures.aap.org/pdfs/Guidelines_PDF/7-Promoting_Physical_Activity.pdf

American Psychiatric, A. (2000). Diagnostic and statistical manual of mental disorders (DSM-IV-TR).

Autism and Developmental Disabilities Monitoring Network Surveillance Year 2008 Principal Investigators, & Centers for Disease Control and Prevention. (2012). Prevalence of autism spectrum disorders--autism and developmental disabilities monitoring network, 14 sites, united states, 2008. *Morbidity and Mortality Weekly Report.Surveillance Summaries (Washington, D.C.: 2002), 61*(3), 1-19.

Bass, M. M., Duchowny, C. A., & Llabre, M. M. (2009). The effect of therapeutic horseback riding on social functioning in children with autism. *Journal of Autism and Developmental Disorders, 39*(9), 1261-1267. doi: 10.1007/s10803-009-0734-3; 10.1007/s10803-009-0734-3

Bhat, A. N., Landa, R. J., & Galloway, J. C. (2011). Current perspectives on motor functioning in infants, children, and adults with autism spectrum disorders. *Physical Therapy, 91*(7), 1116-1129. doi: 10.2522/ptj.20100294; 10.2522/ptj.20100294

Centers for Disease Control and Prevention. (2012a). Basics about childhood obesity. Retrieved November 4, 2012, from http://www.cdc.gov/obesity/childhood/basics.html

Centers for Disease Control and Prevention. (2012b). Facts about ASDs. Retrieved November 4, 2012, from http://www.cdc.gov/ncbddd/autism/facts.html

Fragala-Pinkham, M., Haley, S. M., & O'Neil, M. E. (2008). Group aquatic aerobic exercise for children with disabilities. *Developmental Medicine and Child Neurology, 50*(11), 822-827. doi: 10.1111/j.1469-8749.2008.03086.x; 10.1111/j.1469-8749.2008.03086.x

Krebs, N. F., & Jacobson, M. S. (2003). Prevention of pediatric overweight and obesity. *Pediatrics, 112*(2), 424.

Macdonald, M., Esposito, P., & Ulrich, D. (2011). The physical activity patterns of children with autism. BMC Research Notes, 4, 422. doi: 10.1186/1756-0500-4-422

Nyberg, G. A., Nordenfelt, A. M., Ekelund, U., & Marcus, C. (2009). Physical activity patterns measured by accelerometry in 6- to 10-yr-old children. Medicine and Science in Sports and Exercise, 41(10), 1842-1848. doi: 10.1249/MSS.0b013e3181a48ee6

Oriel, K. N., George, C. L., Peckus, R., & Semon, A. (2011). The effects of aerobic exercise on academic engagement in young children with autism spectrum disorder. *Pediatric Physical Therapy : The Official Publication of the Section on Pediatrics of the American Physical Therapy Association*, *23*(2), 187-193. doi: 10.1097/PEP.0b013e318218f149; 10.1097/PEP.0b013e318218f149

Pan, C. Y., & Frey, G. C. (2006). Physical activity patterns in youth with autism spectrum disorders. *Journal of Autism and Developmental Disorders, 36*(5), 597-606.

Pan, C. Y. (2010). Effects of water exercise swimming program on aquatic skills and social behaviors in children with autism spectrum disorders. *Autism : The International Journal of Research and Practice, 14*(1), 9-28. doi: 10.1177/1362361309339496; 10.1177/1362361309339496

Pan, C. (2011). The efficacy of an aquatic program on physical fitness and aquatic skills in children with and without autism spectrum disorders. *Research in Autism Spectrum Disorders, 5*(1), 657-665. doi: 10.1016/j.rasd.2010.08.001

Pan, C. Y., Tsai, C. L., & Hsieh, K. W. (2011). Physical activity correlates for children with autism spectrum disorders in middle school physical education. Research Quarterly for Exercise and Sport, 82(3), 491-498.

Petrus, C., Adamson, S. R., Block, L., Einarson, S. J., Sharifnejad, M., & Harris, S. R. (2008). Effects of exercise interventions on stereotypic behaviours in children with autism spectrum disorder. *Physiotherapy Canada.Physiotherapie Canada, 60*(2), 134-145. doi: 10.3138/physio.60.2.134; 10.3138/physio.60.2.134

Trost, S. G., McCoy, T. A., Vander Veur, S. S., Mallya, G., Duffy, M. L., & Foster, G. D. (2012). Physical activity patterns of inner-city elementary school children. *Medicine and Science in Sports and Exercise,* doi: 10.1249/MSS.0b013e318275e40b