Effectiveness of Exercise-Based Interventions for Children with Autism: A Systematic Review and Meta-Analysis

Ka-Lam Sam
Hong Kong Baptist University, Hong Kong, China
Email: samsam@life.hkbu.edu.hk

Bik-Chu Chow and Kwok-Keung Tong
Department of Physical Education, Hong Kong Baptist University, Hong Kong, China
Email: {bchow, tongkk}@hkbu.edu.hk

Abstract—Autism is a spectrum of closely-related disorders with a shared core of symptoms. Children on the autism spectrum have problems to some degree with social interaction, communication, and motor coordination. An exclusive exercise program is considered an emerging effectiveness of exercise interventions on exercise mastery, physical fitness, and social competence for children with ASD. Eight studies were employed by a systematic review, then followed by the meta-analysis. Findings indicated that the exercise interventions had a positive effect on developing skills of exercise mastery (d=.57) and social competence (d=.58). Further studies are needed to demonstrate whether early exercise intervention can moderate the sensory processing problems (i.e., poor coordination, under-respond to stimulation) and help regulating emotions of children with ASD.

Index Terms—autism, systematic review, meta-analysis

I. INTRODUCTION

Physical exercise is crucially important for a healthy lifestyle for children with and without disabilities [1], [2]. It has been suggested that children should participate in physical activities that enhance and maintain strength at least twice a week and they should set a goal of accumulating at least 30 min of moderate to vigorous physical activity, preferably every day of the week in order to acquire substantial health benefits [3].

Autism Spectrum Disorder (ASD) is a prevalent neurological symptoms observed even in childhood stages, WHO stated that people with ASD constitute a special risk group because of their sedentary lifestyle, given that this increases the risks of heart disease, diabetes and obesity [4]. Since physical exercise has proven to be an effective means to prevent these problems in the general population, it is likely to also be effective in the ASD population. Pitetti, Rendoff, Grover, and Beets (2007) indeed showed that a walking program not only improved the physical condition but also reduced the Body Mass Index (BMI) of ten adolescents with severe autism [5]. Yilmaz et al. (2004) also found that swimming training is effective for developing physical fitness and water orientation in autistic children [6].

More importantly, as to the assumed beneficial effects on different health domains (motor coordination, balance, flexibility, and etc.) of ASD [7] it was found that, apart from improving balance and flexibility [6] aerobic exercise also reduced the stereotypical behavioral patterns of children with ASD [6], [8], as well as their self-stimulation behaviors [9]. Other studies using exercises as interventions reported positive effects on social behavior [10], [11], communication skills [12], academic engagement [13], and sensory skills [14]. Since then, a holistic search of those exercise intervention studies and quantifying the outcome effects may help to develop an appropriate exercise program for children with ASD, such can be seemed as evidence-based practices indeed.

The interest in the potentially beneficial effects of physical exercise has grown substantially in the last decade. However, there are lack of studies aiming at tidying up those results generated by original intervention studies of physical exercise. Therefore, a global screening of the relevant literature suggests that physical exercise is a sensible approach to addressing a variety of problems associated with ASD. But does the existing literature support this conclusion quantitatively? And can researchers derive more refined recommendations from the findings regarding the type of sport and exercise interventions that are most suitable in the treatment of ASD? To facilitate the evidence-based practice in this area, a systematic review followed by a meta-analysis is worthwhile to conduct.

The purpose of this study was to facilitate evidence-based practice in the area, a systematic review of interventions designed to increase the exercise behavior of children with ASD was conducted initially. The specific aim of this review was to describe the characteristics and exercise interventions of the included...
studies (e.g., participants, how exercise behaviors were taught, and benefits of increased exercise). A review of this type of studies primarily intends to guide and inform practitioners/educators as they develop school and residential programs for individuals with ASD. A secondary aim is to build upon the existing database so as to stimulate future research efforts aimed at using exercise to improve the physical and psychological health of individuals with ASD. To quantify the effect of those exercise interventions, a meta-analysis of quantitative results from independent primary studies on exercise interventions for children with ASD was conducted. The objective of this analysis was to combine intervention/treatment effects into a common metric that is standardized - effect size, which refers to the magnitude of the effect or the strength of the intervention. Meta-analysis itself is an objective and quantitative methodology for synthesizing previous studies or research on a particular topic into an overall finding. An ultimate goal of this study was to identify the most effective exercise intervention for the selection or further development of a quality exercise program for children with ASD.

II. Method

A. Search Procedure

The search covered all electronic databases available, such as MEDLINE/PubMed, EBSCOhost, ProQuest, Scopus, Web of Science, and etc. In the searches the following keywords were used: “Pervasive Developmental Disorders”, “autism”, “ASD/Autism Spectrum Disorder”, “Asperger”, and “PDD-NOS”. The terms with “sport”, “exercise”, “physical exercise”, “physical activity”, “aerobic”, “fitness”, “walk”, “gym”, “jog”, “swim”, “group exercise”, and “individual exercise” were paired. Pertinent articles using the generally accessible websites of ScienceDirect, SpringerLink, SAGE journals online, WILEY online library and Google Scholar were included for further investigation. The Web of Science “related articles” option was exploited to identify any additional articles relevant to our field of interest.

B. Inclusion Criteria

To be included in our meta-analysis, studies have to meet the following criteria: 1) studies needed to have been published between 2004 and 2014; 2) they had to include children with an ASD diagnosis [15]; 3) the interventions described had to involve some kind of physical exercise; and finally, 4) the behavioral effects attributed to the interventions needed to be quantitative such that they could be transformed into percentages reflecting relative behavioral change.

C. Participants

All the participants in the experimental groups from the selected studies must be diagnosed with ASD in accordance with the criteria of the DSM-V [15]. Studies involve control groups were preferred in a higher priority.

D. Data Analysis

To compare the studies on different dimensions, i.e., physiological, psychological aspects, improvement scores of all the potential variables were assessed with the illustration in percentages reflecting the observed behavioral change between measurements. The main concern in this study is the difference between baseline values obtained prior to the exercise or intervention and values obtained shortly after exercise/program completion.

For each individual study, an overall improvement score was calculated. For studies that reported a compound dependent variable consisting of various sub-measures, the total average score of the sub-measures was entered into the analysis. For studies reporting on more than one participant with ASD, the mean percentage score across the participants was calculated.

Those scores were computed as effect sizes, an effect size refers to the magnitude of the effect or the strength of the intervention. In effect size nomenclature, an effect size of \( d = \pm 1.00 \) would suggest a group difference score equal to one full standard deviation between patients in the treatment condition relative to the comparison group, or the difference between pre-post scores. An effect size of \( d = -0.50 \) would suggest a group difference of one half of a standard deviation in the negative, direction. On the other hand, an effect size in the range of \( d = 0.20 \) is considered small, while effect sizes in the range of \( d = 0.50 \) and \( d = 0.80 \) are considered moderate and large, respectively [16].

Effect size (Cohen’s d) of each outcome domain from the respective study was calculated using the equation [16] below:

\[
d = \frac{\bar{x}_i - \bar{x}_c}{\sqrt{(n_i - 1)s_i^2 + (n_c - 1)s_c^2} / (n_i + n_c - 2)}\]

where \( \bar{x}_i \) and \( \bar{x}_c \) are the means of intervention group and control group, \( n_i \) and \( n_c \) are the sample size of the intervention group and control group, \( s_i^2 \) and \( s_c^2 \) are the variance of the intervention group and control group.

The quality of the selected studies was rated, using the Jadad (1996) [17] scale for intervention studies. One point will be given for each of the questions: 1) Was the study described as randomized? 2) Was the study described as double blind? 3) Was there a description of withdrawals and dropouts? Additional 1 point will be given for each of the item: 1) The method of randomization was described in the paper, and that method was appropriate. 2) The method of blinding was described, and it was appropriate. One point will be deducted for each of the item: 1) The method of randomization was described, but was inappropriate. 2) The method of blinding was described, but was inappropriate. Consider the items above, the overall Jadad score for each study was from 0 to 5.

The analysis of the statistics was conducted using two software packages, SPSS Version 21 (www.spss.com) and Comprehensive Meta-Analysis Version 2.2.064 (www.meta-analysis.com).
**TABLE I. SUMMARY OF THE SELECTED STUDIES**

<table>
<thead>
<tr>
<th>Study</th>
<th>N (Male, Female)</th>
<th>Age (M±SD)</th>
<th>Place</th>
<th>Design</th>
<th>Rating</th>
<th>Intervention</th>
<th>Finding</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bahrami (2012)</td>
<td>30 (26, 4)</td>
<td>5-16 (T: 9.20±3.32; C: 9.06±3.33)</td>
<td>Iran</td>
<td>RCT</td>
<td>3</td>
<td>Kata techniques training (56 sessions, 90/min/session; 4 day/week,)</td>
<td>T &gt; C (SS-GARS-2)</td>
</tr>
<tr>
<td>Bass (2009)</td>
<td>34 (29, 5)</td>
<td>4-10 (T: 6.95±1.67; C: 7.75±1.65)</td>
<td>US</td>
<td>RCT</td>
<td>3</td>
<td>Horse riding (1hr/session, 1/session/week; 12 weeks)</td>
<td>N.S. (SP) T &gt; C (SRS)</td>
</tr>
<tr>
<td>Chan (2013)</td>
<td>40 (36, 4)</td>
<td>6-17 (T: 11.28±3.90; C: 12.42±3.25)</td>
<td>HK</td>
<td>RCT</td>
<td>5</td>
<td>Chinese mind-body exercise (1hr/session, 2 session/week; 4 weeks)</td>
<td>N.S. (ATEC-SEN; ATEC-PHY; ATEC-SOC)</td>
</tr>
<tr>
<td>Fragała-Pinkham (2008)</td>
<td>16 (11, 5)</td>
<td>6-12 (9.58±1.33)</td>
<td>US</td>
<td>A-B</td>
<td>1</td>
<td>Aquatic exercise (32-50/min/session; 2 session/week; 14 weeks)</td>
<td>N.S. (M-PEDI; FTS)</td>
</tr>
<tr>
<td>Fragała-Pinkham (2011)</td>
<td>12 (11, 1)</td>
<td>6-12 (T: 9.60±2.60; C: 9.60±1.30)</td>
<td>US</td>
<td>CT</td>
<td>1</td>
<td>Aquatic exercise (40 min/session; 2 session/week; 14 weeks)</td>
<td>N.S. (SCS; IPU)</td>
</tr>
<tr>
<td>Pan (2010)</td>
<td>16 (16, 0)</td>
<td>6-9 (T: 7.27±1.25; C: 7.20±1.89)</td>
<td>Taiwan</td>
<td>CT</td>
<td>3</td>
<td>Aquatic exercise (20 sessions; 90/min/session; 2 session/week)</td>
<td>N.S. (SSBS-2-SC)</td>
</tr>
<tr>
<td>Pan (2011)</td>
<td>15 (15, 0)</td>
<td>7-12 (T: 9.31±1.67; C: 8.75±1.76)</td>
<td>Taiwan</td>
<td>CT</td>
<td>3</td>
<td>Aquatic exercise (28 sessions; 60/min/session; 2 session/week)</td>
<td>T &gt; C (HAAR-IV) N.S. (SnR)</td>
</tr>
<tr>
<td>Pitetti (2007)</td>
<td>10 (6, 4)</td>
<td>14-19 (T: 16.60±1.90; C: 17.40±1.10)</td>
<td>US</td>
<td>CT</td>
<td>1</td>
<td>Treadmill walking (individualized progression program; 9 months)</td>
<td>N.S. (BMI)</td>
</tr>
</tbody>
</table>

**T/C: Treatment group/Control group**

RCT/CT/A-B: Randomized Controlled Trial/Controlled Trial/A-B group
Rating: Based on Jadad scale (0-5) for RCTs
N.S.: Non-Significant difference

SP: Sensory Profile
SRS: Social Responsiveness Scale

**TABLE II. OUTCOME MEASURES ACROSS DOMAINS**

<table>
<thead>
<tr>
<th>Domain</th>
<th>Measure</th>
<th>Description</th>
<th>Study</th>
</tr>
</thead>
<tbody>
<tr>
<td>Exercise Mastery</td>
<td>SP</td>
<td>125-item; by parents/teachers; 5-point Likert scale ranging from 1 (always) to 5 (never)</td>
<td>Bass (2009)</td>
</tr>
<tr>
<td></td>
<td>ATEC-SEN</td>
<td>18-item; by parents/teachers; 1 (not true), 2 (somewhat true), 3 (very true)</td>
<td>Chan (2013)</td>
</tr>
<tr>
<td></td>
<td>M-PEDI</td>
<td>159-item; scale score ranging from 0 to 100</td>
<td>Fragała-Pinkham (2008)</td>
</tr>
<tr>
<td></td>
<td>SCS</td>
<td>16-item; levels ranging 1 to 5</td>
<td>Fragała-Pinkham (2011)</td>
</tr>
<tr>
<td></td>
<td>HAAR-IV</td>
<td>8-item; 0 (unable) to 1 (able)</td>
<td>Pan (2011)</td>
</tr>
<tr>
<td>Physical Fitness</td>
<td>ATEC-PHY</td>
<td>25-item; by parents/teachers; 1 (not true), 2 (somewhat true), 3 (very true)</td>
<td>Chan (2013)</td>
</tr>
<tr>
<td></td>
<td>FTS</td>
<td>Participant is timed while getting up off the floor, walking 3 meters, and returning to a sitting position on the floor</td>
<td>Fragała-Pinkham (2008)</td>
</tr>
<tr>
<td></td>
<td>IPU</td>
<td>A muscle endurance test by isometric push-up (seconds) according to the Brockport Fitness Test manual</td>
<td>Fragała-Pinkham (2011)</td>
</tr>
<tr>
<td></td>
<td>SnR</td>
<td>A flexibility test of the hamstring muscles and lower back from the Taiwan Ministry of Education Physical Fitness Test manual</td>
<td>Pan (2011)</td>
</tr>
<tr>
<td></td>
<td>BMI</td>
<td>Participant’s body mass (kg) divided by the square of their height (m)</td>
<td>Pitetti (2007)</td>
</tr>
<tr>
<td>Social Competence</td>
<td>SS-GARS-2</td>
<td>14-item; by parents/teachers; 0 (never) to 3 (always)</td>
<td>Bahrami (2012)</td>
</tr>
<tr>
<td></td>
<td>SRS</td>
<td>65-item; by parents/teachers; 4-point Likert scale ranging from 0 (never) to 3 (always)</td>
<td>Bass (2009)</td>
</tr>
<tr>
<td></td>
<td>ATEC-SOC</td>
<td>20-item; by parents/teachers; 1 (not true), 2 (somewhat true), 3 (very true)</td>
<td>Chan (2013)</td>
</tr>
<tr>
<td></td>
<td>SSBS-2-SC</td>
<td>32-item; scale score ranging from 32 to 160</td>
<td>Pan (2010)</td>
</tr>
</tbody>
</table>

**T/C: Treatment group/Control group**

RCT/CT/A-B: Randomized Controlled Trial/Controlled Trial/A-B group
Rating: Based on Jadad scale (0-5) for RCTs
N.S.: Non-Significant difference

SP: Sensory Profile
SRS: Social Responsiveness Scale

© 2015 International Journal of Learning and Teaching
III. RESULTS

Fig. 1 shows the summary of the study selection procedure; Table I shows the summary of the selected studies [5], [10], [11], [14], [18]-[21]; outcome measures were categorized into different domains (exercise mastery, physical fitness, and social competence) according to its nature, Table II shows the outcome measures across domains within the selected studies. The effect sizes of each domain are shown as below.

A. Study Selection

Studies retrieved 325 potential articles between 2004 and 2014 from the online research databases that were listed previously. After excluding 76 duplications, the articles were first screened by their title and abstract, resulting in the exclusion of 143. Furthermore, 21 articles were excluded by not fulfilling the remaining selection criteria: not solely ASD subject recruited; no demographic data; no pre and post data; Non-Randomized Controlled Trial/Controlled Trial/A-B (non-RCT/CT/A-B) design; no measurement tools reported; and no baseline measurement. The result of the overall screening left 8 articles for inclusion.

B. Exercise Mastery

Five studies involving 125 subjects (69 treatment, 56 control) evaluated the component of exercise mastery. The treatment conditions were overall superior to those control conditions (d = .57; 95% CI = .22, .92; p < .05; heterogeneity: χ² = 1.49, df = 4, p > .05, I² = 0.00) (see Fig. 2).

C. Physical Fitness

Five studies involving 101 subjects (55 treatment, 46 control) evaluated the component of physical fitness. No significant differences between treatment and control conditions overall (d = .36; 95% CI = -.02, .74; p > .05; heterogeneity: χ² = .53, df = 4, p > .05, I² = 0.00) (see Fig. 3).

D. Social Competence

Four studies involving 119 subjects (62 treatment, 57 control) evaluated the component of social competence. The treatment conditions were overall superior to those control conditions (d = .58; 95% CI = .21, .95; p < .05; heterogeneity: χ² = 2.61, df = 3, p > .05, I² = 0.00) (see Fig. 4).

IV. DISCUSSION

Evidence was found effects on exercise programs for children with ASD. In this study, eight studies were analyzed and the effect sizes of each study were computed. Since the measures among those studies were different in nature, in order to avoid “adding apples and oranges” problem, three groups of outcomes were identified by the consensus of a team of experts in the field of physical education and sports science, and education psychology, these three outcome measure groups are 1) exercise mastery; 2) physical fitness; and 3) social competence. Hence, the results in this study were interpreted separately according to the specific outcome domain it belongs to.
It was expected that outcomes of exercise programs include improved exercise mastery, physical fitness, and social competence [22], even for those children with ASD. However, although the effect size of physical fitness was small to medium (.38), the 95% confidence interval showed that it was not significant (pass though zero). The rest of the two - exercise mastery and social competence, was exhibit a medium to large effect,.57 and .58 respectively.

Why those exercise interventions cannot produce a significant effect on physical fitness? A similar result was shown in Shin and Park (2012)’s study [23], but applied for people with intellectual disability, with no observable/significant changes as well. The possible answers might be: 1) relatively few studies in this area, and those selected studies may not be that representative; 2) small sample size; 3) nature of the selected exercises; and 4) it is the reality.

A highlighted issue of the meta-analysis study is that: with large sample size, contribute most weight to the mean weighted effect size (overall effect size) no matter which model (fixed/random effect model) that it be adopted for.

### A. Sampling Size Estimation for the Follow-up/Future Study

The obtained effect sizes could be used to calculate the require sample size for the similar intervention study as the follow-up/future research. In this study, the overall effect sizes were .57, .36, and .58 for exercise mastery, physical fitness, and social competence respectively; set the alpha level as .05, power as .80, number of groups as 2 (exercise vs control), and number of measurements as 3 (time point: 1-3), the minimum sample size required ranges from 12 to 26 (Table III).

<table>
<thead>
<tr>
<th>Domain</th>
<th>Effect size</th>
<th>α</th>
<th>Number of groups</th>
<th>Estimated sample size</th>
</tr>
</thead>
<tbody>
<tr>
<td>Exercise mastery</td>
<td>.57</td>
<td>.05</td>
<td>.80</td>
<td>2</td>
</tr>
<tr>
<td>Physical fitness</td>
<td>.38</td>
<td>.05</td>
<td>.80</td>
<td>2</td>
</tr>
<tr>
<td>Social competence</td>
<td>.58</td>
<td>.05</td>
<td>.80</td>
<td>2</td>
</tr>
</tbody>
</table>

### B. Assessment of Publication Bias

The publication bias is another issue while people might criticize the validity of the published studies, i.e., positive results may favor for publication. To assess the publication bias, a funnel plot of all eight selected studies was adopted as follows:

![Funnel plot of standard error (SE) by effect size (ES)](image)

Since the funnel plot above was the pooled outcome, including it in the results part might not be that appropriate.

As observed from the funnel plot, those studies formed a symmetrical shape which indicated no sufficient publication bias was found.

### C. Limitation

In this study: 1) non-control-trial studies were excluded; 2) non-English studies were excluded; 3) unpublished articles (conference papers) were excluded; and 4) studies before 2004 were excluded. Those excluded studies might have certain weightings and might change the results of this meta-analysis.

On the other hand, the program duration, frequency, and the length of each session would be the potential factors affecting the overall effectiveness. However, since relatively few studies included in this study compare to Shin and Park (2012) [23], and most of them had similar duration, frequency, and session length, i.e., 12-14 weeks, 2 session/week, and 60-90mins. One study [5] did not report the frequency and session length since it was an individualized progression program. For those reasons, the split-up sub-outcome analysis hence could not be performed.

Furthermore, the nature and the mode of exercise, and the age range of the participants would also be the factors influencing the outcome; as the exercise varies across studies, and the current focus was on children with ASD, but the age range did vary among, from 5 to 19 in these studies, which probably covered some of the adolescences or even adults instead of children. However, those were the solailey available articles from online e-databases with the recent search keywords and screening method, a review of the current searching and screening procedure was recommended in advance.

### V. CONCLUSION

Current research evidence showed that exercise interventions were effective for ASD children in: 1) exercise mastery and 2) social competence. Further studies were needed to demonstrate whether early exercise intervention can moderate the sensory processing problems (i.e., poor coordination, under-respond to stimulation) and help regulating emotions of children with autism.

### REFERENCES


