The University of Rome "Foro Italico"
The German Sport University, Cologne
The University of Southern Denmark, Odense
The Norwegian School of Sport Sciences, Oslo
The University of Vienna

European Master in Health and Physical Activity

(Laurea Magistrale Internazionale in Attività Fisica e Salute) LM 67- I

"The effect of a school-based cluster randomized physical activity intervention study on symptoms of mental health problems amongst 14-year-old students in Akershus county, Norway".

Daniel Inglis
Candidate

Elin Kolle
Supervisor

Runar Barstad Solberg
Co - Supervisor

Academic Year 2018-2020

Abstract

Introduction: The prevalence of mental health problems has shown to increase amongst adolescents. Common interventions are often stigmatized, limiting the likelihood of adolescents seeking help for mental health related problems. Alternatively, physical activity (PA) has been hypothesized to have a positive effect on mental health. However, its evidence-based practice has proven to be scarce. This thesis evaluated the effect of a 9-month long school-based PA intervention on total mental health problems among Norwegian adolescents. Further, evaluating intervention effect differences in respect to adolescent's mental health problems at baseline.

Methods: Eleven secondary schools (N = 886 students, 14 y) was cluster-randomized into three groups; Intervention model M1 (n=3), intervention model M2 (n=4) and control (n=4), with the intervention models implementing additional 120 min of PA, each week. General mental health problems were assessed by a strength and difficulties questionnaire (SDQ). A linear mixed model was used to evaluate the intervention effect on SDQ scores, and differences between the effect in relation to the degree of SDQ scores at baseline.

Results: No significant intervention effect was seen in total SDQ scores. Nor did the effect vary significantly in relation to the degree of SDQ scores at baseline, concerning both interventions and genders, compared to control. A positive effect was seen in the subcategory of "prosocial behaviors" and "relationship with peer problems" amongst boys in the M1 intervention (P = 0.024 and P = 0.045), and a negative effect on "relationship with peer problems" amongst girls in the M2 intervention (P = 0.012).

Conclusion: The two interventions did not provide any significant effect on total mental health problems. However, implementing one additional PE class, alongside 30 min of free activities, and PA-based learning (M1), may enhance social relationships amongst boys. Conversely, additional PA through freely chosen activities, little structure and supervision by teachers (M2) may negatively affect peer relationships amongst girls.

Acknowledgments

With this thesis, i will have completed a two year long european masterprogram, which have offered me much more than just a degree. I feel lucky to have gotten the opportunity to learn from a wide range of scientific leaders within their respective fields. Throughout my degree, both national and international friendships have developed, in which have allowed for shared experiences concerning cultures and with this, an international perspective on the field of health and physical activity.

In regards to this thesis, i would like to acknowledge my supervisor Elin Kolle whom provided great guidance and feedback for the development of this thesis. Also, my co-supervisor Runar Barstad Solberg have had an essential part of me being able to successfully complete this thesis. With his knowledge and great attitude, he was able to pick me up and kept me going under this challenging period. With this, I thank you both.

I would also like to extend my greatest appreciation to my family, my girlfriend, and my close friends. You have all shown me an immense support throughout this journey. I could never thank you enough.

Table of content

Abbreviations

ScIM = School in Motion

PA = Physical activity

PE = Physical education

MVPA = Moderate to vigorous physical activity

SDQ = Strength and difficulties questionnaire

BDI = Beck Depression Inventory

DADS = Duke Anxiety-Depression Scale

HADS = Hospital Anxiety-Depression Scale

HSCL = Hopkins Symptom checklist

CBCL = Child behavior check list

YSR = Youth Self Report

SDT = Self determination Theory

PAQ-C = Physical Activity Questionnaire for children

1. Introduction

Mental health has shown to decline in recent years, and disorders related to poor mental health has seen an increase amongst adolescents. Difficulties related to mental health is one of the leading causes of the overall disease burden worldwide, and years lost due to disability, with anxiety and depression disorders being the biggest contributors amongst the adolescent population (Vos et al., 2015; Ritchie & Roser, 2018; Pascoe et al., 2020).

The onset of mental health disorders and poor mental health amongst the population can often be traced back to the adolescent years. However, this with only a few cases ever being reported (World Health Organization, 2019). Interventional treatments for poor mental health and its disorders such as psychotherapy and pharmaceutical methods is often stigmatized in a negative manner. Consequently, adolescents whom experience mental health related problems are less likely to seek help, which in turn may affect their quality of life negatively (Pascoe et al., 2020). Having recognised the growing issue, the literature has expressed the need for viable interventions that may be more suitable for this target population, in terms of supporting their mental health and prevent its disorders with less stigma attached (Pascoe et al., 2020).

As an alternative method, physical activity (PA) and exercise have previously been hypothesized as a more self-sustainable and cost-effective interventional method, in which to decrease mental health problems, and prevent its onset, and with less stigma associated with its practice (Pascoe et al., 2020; Liu, Wu & Ming, 2015). Numerous studies have reported a favorable effect on various mental health variables, following PA interventions in adolescence (Biddle & Asare, 2011; Biddle, Ciaccioni, Thomas & Vergeer, 2019). However, despite showing promising interventional effects on mental health, its evidenced based practice concerning mental health problems has proven to be scarce, and more studies are needed in order to support its findings (Brown, Pearson, Braithwaite, Brown & Biddle, 2013; Biddle et al., 2019).

To address the lack of literature, the aim of this thesis was to investigate the effect of a multicomponent school-based PA intervention with two different intervention models,

on total mental health problem scores, amongst 14-years old students from Akershus county, Norway.

1.1 Research question

What is the effect of a 9-month long school-based PA intervention, on subjectively reported total mental health problem scores among Norwegian 14-year old adolescents? Further, how does the intervention effect vary in respect to the adolescent's mental health score at baseline?

The following two hypotheses will be tested in this thesis.

Total mental health outcome

H0: The two PA intervention models did not have a significant effect on the student´s total mental health problem score when compared with the control group.
H1: The two PA intervention models did have a significant effect on the student´s total mental health problem score when compared with the control group.

H0: The intervention effect did not vary in respect to the student´s total mental health problem score reported at baseline.
H1: The intervention effect did vary in respect to the student´s total mental health problem score reported at baseline.

2.0 Theory

2.1 Mental health amongst adolescents.

Mental health can be defined as a state of well-being, supplemented with an ability to overcome and cope with a variety of stressors in living, through the realization of one's own abilities. It involves a feeling of personal value, capacity to successfully contribute to society and to be able to build relationships with peers (World Health Organizaion, 2004; Keyes & Ryff, 2002).

The previous concepts of mental health often relate back to the two ancient Greek approaches; 1) the hedonic and 2) the eudemonic tradition. According to the hedonic tradition, mental health is achieved through positive emotions, by minimizing pain and maximizing pleasure in everyday living. Conversely, in the eudemonic tradition mental health is rather viewed as the end result from optimal functioning and overcoming obstacles through personal effort. Besides having two different views on mental health and its existence, both the hedonic and eudemonic tradition often coexist in modern literature, emphasizing the importance of resilience and seeking positive experiences in everyday living (Guo, Tomson, Keller & Söderqvist, 2018).

2.1.1 The self-determination theory

As a more modern approach to mental health, the self-determination theory (SDT) explains how the social environment surrounding an adolescent may impact his or her well-being and functioning, by fulfilling the following three basic psychological needs;

- Autonomy: The feeling of acting in concordance with one´s own choice and behavior.
- Relatedness: A sense of belonging and being able to socially connect with individuals belonging to one's environment
- Competence: Feeling of mastery through successfully interacting with one´s social environment and by expressing one´s capabilities (Doré et al., 2020).

Stepping into adolescence is known to be a challenging period and involve considerable changes in the neurological and physical domain (Herpertz-Dahlmann, Bühren & Remschmidt, 2013). On top of these internal changes, psychosocial factors including relationships to both family members and friends are often seen challenged (Herpertz-Dahlmann, Bühren & Remschmidt, 2013). As a consequence of this, an adolescent may experience a shift in their social environment, which could put additional strains on mental health (Ueno, 2005). During this transitional phase, positive psychological factors and cooping abilities has proven to act as important mediators in mental health amongst adolescents. To elaborate, research show that adolescents whom possess a greater ability to overcome adversity, may be more resilient to mental health problems and the development of its disorders (Steinhausen & Metzke, 2001).

2.2 Mental health disorders amongst adolescents

Mental health disorders and problems are often associated with abnormal emotions, thoughts, behaviors, lack of self-esteem and dysfunctional relationships with peers (WHO, 2020a; Ritchie & Roser, 2020). External factors such as an adolescents social and economic circumstances have also shown to play a role in the onset of poor mental health and disorders, including poor financial status amd lack of social relationships with peers (WHO, 2020b). Consequently, mental disorders overall comprise of a broad spectrum of symptoms and problems, in which may vary amongst the different disorders and makes it difficult to generalize.

Emotional disorders have been reported to be one of the more prevalent contributors to poor mental health amongst adolescents, with depression and anxiety being the respectively fourth and ninth leading cause of disability amongst adolescence at the age of 15-19 years (WHO, 2020b). In fact, studies show that approximately 50% of mental health disorders amongst adults may have had its onset prior to 18 years of age (Das et al., 2016). Mental health problems and disorders, including depression and anxiety, has shown to be more prevalent amongst boys during childhood, compared to girls of the same age group (Derdikman., 2012). However, evidence show that upon reaching the age of 15 years, the prevalence of mental health disorders and poor mental health, increase amongst girls, further exceeding the prevalence amongst boys. Because of this,

poor mental health and its disorders are often more prevalent and greatly reported in adolescent girls (Derdikman., 2012). The gender distribution of mental health disorders and problems has previously been hypothesized to differ between the genders in respect to hormonal changes, brain development, genetical factors and environmental differences (Derdikman., 2012).

In general, adolescents may often experience poor mental health, despite not being diagnosed with mental health disorders. As a result, the literature emphasize the importance of continuing promoting positive mental health, as well as working towards preventing symptoms of mental health disorders amongst the adolescent population. Overall, this emphasizes the importance of further investigating methods that may positively influence mental health and symptoms of mental health problems effectively (Das et al., 2016).

2.3 Measuring mental health problems amongst adolescents.

Assessing mental health related problems amongst the adolescent population is commonly performed using self-reported psychometric questionnaires (Deighton et al., 2014; Williams, Noël, Cordes, Ramirez & Pignone, 2002). In practice, various questionnaires have been constructed to evaluate different aspects of mental health in adolescents. Some psychometric measurements aim to assess explicit symptomatology, including symptoms related to depression and anxiety disorders. Examples include the Beck Depression Inventory (BDI), Duke Anxiety-Depression Scale (DADS), and Hopkins Symptom Checklist (HSCL) questionnaires (Williams, Noël, Cordes, Ramirez & Pignone, 2002). As being specified to measure symptoms related to a specific mental health disorder, the questionnaires are known to provide more detailed measures and with a greater ability to detect changes over time. However, as these questionnaires are often limited to specific symptoms and their respective diagnostics, the results may be difficult to compare across literature, as well as various cases of mental health problems amongst adolescents (Deighton et al., 2014).

As a more general measure, a handful of questionnaires have been designed with a purpose to evaluate mental health in a much broader spectrum (Deighton et al., 2014).

These general questionnaires measure mental health by symptoms commonly associated with mental health problems amongst adolescents overall. To give some examples, the Beck Youth Inventories (BYI), Behavior Assessment System for Children (BASC), Strength and Difficulties Questionnaire (SDQ), Achenbach System of Empirically Based Assessment (ASEBA), Youth Outcomes Questionnaire (YOQ), and Kidscreen are all commonly used questionnaires with this purpose (Deighton et al., 2014).

Amongst these, the SDQ has been sought as one of the more popular questionnaires by researchers and aims to evaluate symptoms of mental health problems (Duinhof et al., 2020). The reason for its popularity stems from its simple structure, good validity in terms of assessing psychopathological symptoms in adolescence, and its sufficient reliability. The questionnaire is also considered to be brief in its nature, in respect to comparable questionnaires (Essau et al., 2012; Deighton et al., 2014). Due to this, the SDQ is known to be applicable when measuring general symptoms of mental health problems amongst the adolescent population, as applied for in this current thesis (Essau et al., 2012; Duinhof et al., 2020; Deighton et al., 2014).

2.3.1 Strength and Difficulties Questionnaire

The SDQ can be administrated to parents or teachers when assessing behaviors in children, and as a self-reporting practice for adolescents between the age of 11-16 years. The questionnaire has a purpose of measuring various factors related to an adolescent's behavior, social relationships, and emotions. Several versions of the questionnaire have been modified to suit different groups and their respective criteria. The most used version of SDQ consist of 25 attributes, aimed to report both negative and positive behavioral traits. In its structure, the attributes can be fitted in to 5 subscales; emotional symptoms, conduct problems, hyperactivity and peer relationship problems. The fifth and last subscale aims to evaluate prosocial behaviors, which is a positive behavioural trait. Therefore, this specific subscale does not count towards the total make up of the mental health problem score (Goodman, 1997; Goodman & Scott, 1999).

As a self-reported measure, the SDQ has previously provided valid results in terms of being able to assess mental health problems amongst adolescents in different countries (Duinhof et al., 2020; Goodman., 2001; Hawes & Dadds., 2004; Koskelainen,

Sourander & Vauras, 2001; Rønning, Handegaard, Sourander & Mørch, 2004). As one example, Essau et al (2012) evaluated the validity of the self-reported version of the SDQ, across five different European countries, including Sweden, Germany, England, Italy and Cyprus. Conclusively, the results indicated that all four translations of the SDQ, including the standard English version, provided valid and consistent results in a majority of the countries. It was further reported that in some instances, the reversed items included in the SDQ, did pose some difficulties when translated into another language, suggesting that these items may be re-worded to allow a more valid measure and comparisons of SDQ scores between countries (Essau et al., 2012). Similarly, Duinhof et al (2020) also reported to find a better comparability of the SDQ total difficulty score when removing the reversed items from the questionnaire (Duinhof et al., 2020).

The validity of the Norwegian version of the self-reported SDQ has also been previously tested amongst Norwegian adolescents. As reported by Richter, Sagatun, Heyerdahl, Oppedal & Røysamb (2011), the Norwegian version of the self-reported SDQ was able to measure symptoms of mental health problems amongst Norwegian adolescents with a reasonable validity. Having said that, the authors raised some issues by analyzing some of the consisting subscales, further suggesting the use of the total difficulty score derived from the SDQ more heavily, opposed to scores derived from individual subscales (Ritcher et al., 2011).

When compared to other general mental health questionnaires, the SDQ proves to adequately assess both positive and negative behaviors, with the ability to differentiate between mental disorder and normative behaviors (Goodman & Scott., 1999). For instance, Goodman & Scott (1999) reported to find comparable results concerning mental health between the SDQ and the Child Behavior Checklist (CBCL), when completed by parents. Similar results were also reported by Klasen et al (2000), whom compared the German version of the SDQ, up against the CBCL and the Youth Self Report (YSR) (Klasen et al., 2000).

2.4 Physical activity

Physical activity is defined as any movements of the body, produced by forces generated by skeletal muscles in which result in energy expenditure (Caspersen, Powell & Christenson, 1985). It is commonly defined by frequency, duration, type and intensity of the activity itself. PA can be performed in many forms and contexts such as regular exercise, with organized sport, free play, and curicullar based activities at school (Green & Smith, 2016).

2.4.1 Measurements of physical activity

Measuring PA can be performed by a direct or an indirect method. Self-reported questionnaires, activity diaries, direct observation, accelerometers, pedometers and heart rate monitoring are some of the more commonly used methods. Each method has its own strengths and weaknesses, which limits their application to the aim of the research, available resources, and target population (Fairclough & Lubans, 2020; Piggin, Mansfield & Weed, 2017).

1) Self-reported measures

Self-reported measurements have traditionally been one of the most common method used when measuring PA behavior on a population-based level (Piggin, Mansfield, & Weed, 2017). These methods are known be less costly, more feasible, and less intrusive in its practice, and therefore often more accepted when used in a larger population. The most commonly used measurements involves the use of logs, diaries and questionnaires (Fairclough & Lubans, 2020).

PA questionnaires can often be categorized in to three different classifications; short-term recall questionnaires, global questionnaires and quantitative history recall questionnaires. Short-term recall questionnaires often involve 7-20 questions related to participant´s PA behavior from the past 7 days and up to a month. In comparison, the global questionnaires are often shorter and consist of 1-4 questions and is therefore

often sought as a feasible tool when measuring PA on a global scale. However, the questionnaire is only able to register little information related to the participant´s PA behavior, due to its brevity. Lastly, the quantitative history recall questionnaires require participants to recall their PA behaviors from the past year, sometimes even longer. The questionnaire may consist of as many as 60 PA behavior related questions, providing additional information concerning behaviors, compared to the two previously mentioned questionnaires.
Having said that, the quantitative recall questionnaire has shown to pose challenges in terms of participants not being able to validly recall past PA behaviors (Piggin, Mansfield & Weed, 2017).

As reported in previous studies, the use of subjective methods to accurately estimate PA behaviors amongst adolescents may provide inconsistent results (Kavanaugh, Moore, Hibbett & Kaczynski, 2015). For instance, Kavanaugh, Moore, Hibbett & Kaczynski (2015) reported to find a poor correlation between subjectively measured moderate to vigorous physical activity (MVPA) by the physical activity questionnaire for older children (PAQ-C), and the MVPA volume measured by accelerometry, amongst young adolescents. Consequently, the author emphasized the importance of selecting validated outcome measures, when aiming to measure MVPA amongst this population (Kavanaugh, Moore, Hibbett & Kaczynski, 2015).

2) Objective measurements

Compared with subjective measurements, objective measures of PA are sought to as a more precise method when measuring PA behaviors. In research, objective measurements often include the use of direct observation techniques, pedometers, heart rate monitors and accelerometers (Piggin, Mansfield & Weed, 2017).

Direct observation is a technique often used in smaller studies, as it allows researchers to make direct observations of PA in a free-living situation. This method allows researchers to obtain additional information regarding influencing factors such as environmental factors and PA context, which may be of value when evaluating PA behaviors (Piggin, Mansfield, & Weed, 2017). Due to this, direct observation is often recognized as the criterion standard, when measuring physical activity amongst

adolescents (Sirard & Pate, 2001). Despite its recognition, there are some disadvantages reported with its practice, which includes subject reactivity where observers may impact PA behaviors with his or her presence. Additionally, direct observation as a research practice is known to be costly and time consuming, limiting it to research purposes entailing a smaller population pool (Fairclough & Lubans, 2020).

Physical activity can be objectively measured with the use of a pedometer. A pedometer is a small sensor that can be worn on one's body, tracking the number of steps performed over a period of time. To further acquire information regarding physiological responses during PA, researchers often implement heart rate monitors in conjunction with other movement-based tracking devices (Brage et al., 2006). With a heart rate monitor, information regarding frequency, duration and intensity of a PA bout can be observed by estimating an individual´s heart rate (beats per minute), and further calculate one's energy expenditure (Fairclough & Lubans, 2020).

Lastly, accelerometers are small portable devices that can be worn on the waist, hip, arm, wrist and the ankle. An accelerometer measures PA by recording changes in acceleration across three orthogonal planes, further expressed as the amount of counts over a specific duration of time, also known as an epoch. With the use of an accelerometer, researchers may be able to record the intensity, frequency and duration of PA and moreover estimate the amount of energy expended during the PA duration (Bouten, Koekkoek, Verduin, Kodde & Janssen. 1997). Furthermore, due to its well investigated validity and its portable size, the accelerometer is known to be one of the more suitable instruments when aiming to measure free-living PA (Piggin, Mansfield & Weed, 2017).

2.5 Physical activity and mental health: Possible mechanisms of its effect

Positive mental health effects following exercise and involvement in PA and sport, have been hypothesized to be caused by several factors and mechanisms, in which may all collectively contribute to prevent and aid symptoms of mental health problems and disorders amongst adolescents. These factors and mechanisms include psychosocial

mechanisms, neurobiological and behavioral changes, in which will be explained in greater detail (Doré et al., 2020; Lubans et al., 2016).

Psychosocial mechanisms

Participating in PA and sport can support mental health amongst adolescents by satisfying basic psychological needs. This includes the feeling of relatedness by interacting with team members, self-efficacy through mastering acitivty related skills, and autonomy by participating independently (Lubans et al., 2016). PA may also provide an adolescent with a sense of importance when acting in his or her respective roles, and with its success, social connections with others involved in similar activities may be further facilitated (Doré et al., 2020).

Neurobiological changes and mental health

PA has been hypothesized to facilitate for changes to the structural composition of the brain and by this, enhancing its functioning and general mental health. Furthermore, exercise and PA have been known to provoke the release of endogenous opioids which is widely recognized in our society as a positive effect following exercise. Among opioids, the release of endorphins and its relationship to the feeling of well-being and euphoria, is often discussed in literature (Lubans et al., 2016). Nevertheless, there is little evidence that support this acute euphoric feeling to produce any long-term effect on adolescent's mental health. New evidence, however, hypothesize that this feeling is rather contributed by several brain monoamines, including noradrenaline, dopamine and serotonin (Lubans et al., 2016).

Behavioral changes and mental health

Performing PA may further improve sleep amongst adolescents, imposed by its exhausting tendencies. Several studies have reported improvements in sleep behaviours amongst adolescents, including its efficiency, onset of sleep latency as well as an improvement in sleep duration. Following this effect, adolescents may develop a healthier sleep pattern, in which could be a moderating factor for mental health overall (Lubans et al., 2016).

As mentioned earlier, self-regulation and the ability to overcome obstacles have shown to be an important factor in mental health and in preventing the onset of disorders

amongst adolescents. When performing PA and sport, adolescents may experience physical and mental obstacles through its practice. By successfully overcoming these challenges, PA may support their ability to overcome these obstacles, potentially making them more resilient to future challenges to come (Lubans et al., 2016).

2.6 Physical activity and mental health in adolescence

In recent years, there has been numerous studies investigating PA and its effect on mental health in adolescents, with several reviews analyzing the associations between two factors (Biddle, Ciaccioni, Thomas & Vergeer, 2019). In a review of reviews Biddle & Asare (2011) concluded the association between PA and mental health in children and adolescents to be indisputable. However, reviewed studies have shown to define and assess both positive mental health, PA, and symptoms of mental health problems in various ways, resulting in inconsistent measures (Biddle & Asare., 2011). When disregarding mental health and limiting the outcome to mental health problems only, the literature has reported various results concerning the effect and the association between PA and mental health problems. Amongst these, depression is known to be one of the more popular topics of research in adolescents (Rodriguez-Ayllon et al., 2019).

Across ten systematic reviews, Biddle et al (2019) reported to find inconsistent results, ranging from a null to a small association between PA and depression amongst adolescents. Similarly, the intervention effect was seen to vary, ranging from an effect size of -0.41 and -0.61 across six meta-analyses. Amongst interventions the effect of PA has shown somewhat stronger results in adolescents whom were diagnosed, and with greater symptoms of depression, compared to adolescents whom are of healthy population. Due to this, the literature suggests that PA may have a more prominent effect amongst adolescents whom possess greater overall symptoms. Despite the topics growing popularity, the effect of PA on depression and mental health problems overall, have not yet been concluded in adolescents. As a consequence, the literature has emphasized the need for further studies with a greater methodological validity to support current findings (Biddle et al., 2019).

Physical activity and Symptoms of mental health problems by SDQ in adolescents

Expanding the outcome of mental health to a more general measure ,symptoms of mental health problems assessed by the self-reported SDQ has shown a low to moderate association with PA behaviors across most cross-sectional and longitudinal studies in adolescents (Kuiper, Broer & Van, 2018; Vedøy et al., 2020; Sagatun et al., 2007; Bell et al., 2019; Wiles et al., 2008; Hallal et al., 2015).

In a cross-sectional study including 96 617 Dutch students (12-17 years of age), Kuiper and colleagues (2018) observed a significantly greater weighted average SDQ score in adolescents who self-reported to be inactive, compared to their active peers ($\beta = 1.12$; 95% CI: 1.10-1.14; $P<0.001$). It was also reported that inactive adolescents possessed a significantly greater score in peer problems, problems with conduct, emotional problems, hyperactivity, and a lower level of pro-social behavior (Kuiper, Broer & Van, 2018). However, in a sample of 599 Norwegian 13-year old adolescents Vedøy and colleagues (2020) reported no association between accelerometer measured PA and student's total weighted mental health problem score, measured with SDQ ($\beta = 0.06$; 95% CI: -0.23-0.36; $P=0.677$) (Vedøy et al., 2020).

Similarly, investigating prospectively, several longitudinal studies have supported no significant association between PA and total symptoms of mental health problems scored by SDQ specifically (Sagatun, Søgaard, Bjertness, Selmer & Heyerdahl, 2007; Wiles et al., 2008; Hallal et al., 2015; Bell et al., 2019). However, some significant associations have been reported in relation to specific subcategories.

Sagatun, Søgaard, Bjertness, Selmer & Heyerdahl (2007) investigated the association between the amount of self-reported PA performed on a weekly basis, and corresponding mental health problem score measured by SDQ, three years later amongst adolescents (15-16 years of age). Conclusively, the study found no significant association between self-reported PA volume and total symptoms of mental health problems score among neither gender. A significant association was however found between the numbers of hours spent performing PA at the age of 15-16 years and emotional symptoms ($\beta = -0.009$; 95% CI: -0.15, -0.03 $P=0.020$) at the age of 18-19

years among boys. Whereas no association was found among girls following the adjustment of confounding factors (Sagatun et al., 2007).

Agreeingly, Wiles and colleagues (2008) reported no association between self-reported PA volume and total SDQ score. However, a lower mean score in emotional problems was observed amongst adolescents whom met the recommended 60 minutes of PA per day, compared with the ones who did not (n=1446, age=11-14 years of age). The mean difference was reduced when adjusting for gender and baseline emotional symptoms (-0.29, 95% CI -0.60, 0.018) (Wiles et al., 2008).

Similar findings were also revealed by Hallal and colleagues (2015) whom at first reported a significant inverse relationship between children (11 years of age) whom adhered to the recommended amount of self-reported PA (300 min/wk) and a lower score of symptoms of mental health problems, four years later at the age of 15 (β = -0.744 95% CI -1.363, -0.125; P=0.04), compared with adolescents who did not. However, when adjusting for confounding factors the association proved to be significant for boys only (β = -1.237 95% CI -2.218, -0.255; P =0.04) (Hallal et al., 2015).

Bell and colleagues (2019) investigated the prospective association between objectively measured PA by accelerometry, and symptoms of mental health difficulties measured with SDQ in a sample of 928 students (12-13 years of age). Follow-up test was performed three years after baseline. Using a more accurate measure of PA, the analysis showed no association between PA, mental wellbeing or symptoms of mental health problems by the SDQ. A significant inverse association was reported between PA volume at the age of 12-13 years and emotional symptoms at the age of 15-16 years (β = -0.11; 95% CI: -0.23, -0.00; P<0.04), however, the association attenuated and was no longer significant after adjusting for gender (Bell et al., 2019).

School-based PA interventions on mental health problem outcomes in adolescents

The effect of PA interventions has previously been evaluated in schools with the aim of promoting numerous health behaviors amongst its students, with various results (Hills et al., 2015; Theron, Liebenberg & Malindi, 2014).

Nevertheless, few studies have investigated the effect of a school-based PA program on mental health problems amongst the adolescent population (Eather et al., 2015).

Amongst few, Bonhauser and colleagues (2005) implemented a total of three PA sessions as curricular practice on a weekly basis (90 minutes per session), lasting a whole school year. The aim was to evaluate the effect of school-based PA intervention on mental health outcomes, including depression and anxiety symptoms, amongst 198 adolescents (n = 98 intervention, n = 100 control) (15 years of age) from a poor socioeconomic background, in Chile. As a result, the study reported to find a significant intervention effect in which reduced anxiety symptoms amongst adolescence in the intervention classrooms, compared to control classrooms ($P < 0.01$). However, no significant effect was seen in depression scores, where both emotional symptoms were measured with a hospital anxiety and depression scale (HADS) (Bonhaiser et al., 2005).

School-based PA interventions and the outcome of general mental health problems by SDQ in adolescents

Eather (2016) examined the effect of an eight week-long, high intensity strength-based exercise regime, on mental health outcomes (measured with SDQ and Physical Self-Description Questionnaire) amongst 15-year-old students (N = 95, N= 51 intervention group, N=45 control group). The intervention was implemented as a part of physical education (PE) classes (60 minutes) and sport classes (60 minutes), equivalating to a total of 120 minutes per week. No significant intervention effects were observed on the total mental health difficulty score measured by SDQ, amongst the whole interventional group when compared to the control. However, the authors reported that students whom were categorized as being at risk, meaning that they scored above 15 on the SDQ at baseline (n=8), benefited more greatly from the intervention compared to adolescents whom were not at risk. However, these findings were not significant (Eather et al. 2016).

In a similar study on children, Bunketorp, Malmgren, Olsson, Linden & Nilsson (2015) investigated the effect of a school-based PA intervention on different variables including mental health problems measured with SDQ. The controlled quasi-experimental study included pupils from 0-6th grade belonging to an interventional

school (n=182) and a control school (n=167). In the interventional school, two additional PA based classes were introduced for one semester, as part of the PE curriculum, lasting around 40 mintes at a time. The two additional PA classes involved activities that supported playfulness and physical motion through activities, emphasizing noncompetitive play and enjoyment through activity and sports. Upon finishing the intervention period, no significant difference in total SDQ score between the control and interventional groups across all grades (0-6th grade) were found (Bunketorp et al., 2015).

With a short summary concerning both the association and the effect of PA on mental health problems amongst adolescents, it is evident that the literature may lack sufficient evidence in order to conclude its effects. Furthermore, this emphasizes the need for randomized controlled school-based interventions examining the effect of PA on mental health, and especially its effect on the outcome of general symptoms of mental health problems, among adolescents.

3.0 Methods

The School in Motion (ScIM) project was conducted as a school-based cluster randomized controlled trial, with three intervention arms and was completed during the school year 2017-2018. The main aim of the ScIM study was to positively affect the students PA levels by implementing additional curiccular based activities. The Norwegian School of Sport Science (NSSS) acted as the main institution of coordination, with three collaborating study partners. Thirty schools from municipalities around the four test centers was included in the study. In the inclusion process, the population density was considered. Exclusion criteria was private and special schools, schools with fewer than 25 students in 9th grade, in addition to schools that participated in other projects with increased physical activity in the school curriculum.

The study sample for this thesis is a subgroup recruited from schools located around the geographical area of the NSSS institution, in Akershus county, Norway.

3.1 Sample population

A total of 22 high schools belonging to Akershus county was invited to participate in the ScIM study. Correspondingly, 11 schools declined, and 11 schools accepted the invitation and agreed to participate in the study. Each school were randomly assigned into either of the three intervention arms, distributed by a neutral third party. One of the control schools withdrew after the randomization process but before baseline testing was completed, leaving four schools in the control group, three schools in intervention model M1 and four schools in model M2. Overall, 1149 students in 9th grade were invited to participate in the study, in which 886 provided a signed consent (77% participation rate). There was one school that withdrew from the project before its completion. This was after finishing a total of twelve weeks of the M2 intervention. Consequently, data from this specific school included values from baseline only, with no excisting follow-up measures.

3.2 Intervention models

The intervention period lasted from September 2017 to June 2018, and participation in activities was mandatory for all students involved in the study. The main aim of the ScIM project was to increase the students PA levels by implementing an additional 120 minutes of PA, on top of their existing 120 minutes of PA from their curricular PE classes.

When developed, the interventions of the ScIM project was constructed on the basis of a social-ecologic understanding, recognizing the environmental and personal influences on one's behavior, and how they relate with one another. In the ScIM study, two different intervention models (M1 and M2) were developed and tested during a 5-month pilot study before the main intervention. The first intervention M1 (Physical active based learning) was developed and implemented on the basis of previous interventions and its effect when implemented amongst children in primary schools. The M2 ("Don't Worry, Be Happy") intervention, however, had little experience behind its implementation, where the intervention had an equal focus on enhancing PA and facilitate for social development amongst the adolescents. Participating schools were allocated additional means from the Directorate of Education, based on the number of pupils at each school, funding one extra hour of PA in their weekly based curriculum. In addition to this, the study requested participating schools to devote 5% of their current weekly curricular schedule (1 hour), to make room for additional PA classes. Consequently, schools participating in the two intervention models had to perform a total of two additional hours of PA, on a weekly basis. Schools belonging to the control group was instructed to adhere to their already excisting PE curriculum, meaning that no additional PA classes were implemented.

Interventional model 1: Physical activity-based learning (M1)

The M1 intervention consisted of three of the following components:

1. Physical activity-based learning: PA was implemented as an educational tool into curricular practice, where various subjects would be taught by including

components of PA. This component were to be implemented by teachers for a total of 30 minutes, each week.

2. Extra PE class: one additional PE class (60/45 minutes) was implemented into the existing standard curriculum. The extra PE class focused on implementing activities with high intensities, and that adhered to student's PA interests.

3. Physical activity: Students were to perform an additional 30 minutes of PA each week, which was not related to educational purposes, and focused on enjoyment through activities. These activities were to be constructed to suit the student's interests.

Interventional model M2: Don't worry - Be happy

The schools belonging to interventional model M2, implemented one extra hour of PE class, referred to as the "Don't worry"-class, and one additional hour of movement-based activities, referred to as the "Be happy"-class. However, schools could freely choose a name for their two interventional hours as they would see fit.

The "Be happy"-class aimed to develop friendships and to allow adolescents to socialize with one another based on mutual interest of activites. The component allowed the adolescents to group up with peers across classrooms, meaning that groups could be formed ourside of classroom allocations. With help from their PE teachers, groups would choose and commence in activities that was of personal interest and value for the students in each group. Furthermore, each individual groups were instructed to develop activity related goals and a plan in which to reach these goals throughout the intervention period. As a general aim, "Be happy"-classes required groups to be more self-organized and independent, with little supervision by teachers. In the "Don't worry"-class, students would go back to their original classroom allocations and continue their activities, related to the " Be happy" class, with a PE teacher present for assistance if needed.

3.3 Intervention adherence

With electronic registration, participating schools were asked to report the amount of time spent on each component of interventional model 1 and model 2. All registrations were assessed weekly by a project coordinator at the NSSS. Total adherence rate was calculated starting from week 38 (2017), and to the end of week 19 (2018). Several weeks were excluded from the equation, due to them being school holidays. This included Week 40, 41, 51, 52 (2017) and week 7, 8 and 13 (2018). In total, completion rate of activities was calculated based on a total of 29 weeks.

3.4 Ethics

The study was approved by the Norwegian Centre for Research Data and adhered to the declaration of Helsinki. Parents and students were asked to give their written consent prior to participation in the ScIM study and was informed that this consent could be revoked by the parents or adolescents at any time.

3.5 Power calculation

In the ScIM study, the main aim was to estimate the effect of implementing additional curiccular based activites, on overall physical activity levels in students. Due to the aim, the design was set to identify a seven percent difference (49 counts per minutes) in total PA levels, between students in the control and the two intervention groups. For the power calculation, the following assumptions were made: 150 cpm in standard deviation (SD), with a power of 90%, and with a significance level of 0.05. As a result, a total of 492 students were required for each group, with an additional 98 students to allow for a 20% loss due to follow-up (total of 590 students). Consequently, this would accommodate to ten clusters for each individual study arm.

The power calculation in the ScIM study was only applied for the effect of additional PE and PA on PA levels in students and not for total mental health score by the SDQ, which stands as the aim of this study. Therefore, a power calculation was not performed for this current thesis.

3.6 Data collection and measurements

Students were tested and all data was collected prior to the start of the intervention and again at the end of the interventional period. Each test was conducted by a team of experienced testers including staff members, PhD candidates and undergraduate students belonging to the NSSS. All anthropometrical measurements including; height (cm), weight (kg), and waist circumference (cm) was performed at each participating school. This was followed by a self-reported strength and difficulties questionnaire in which was administered as a part of an overall health questionnaire, by the use of a computer software (SurveyXact).

Anthropometrical measurements

Height was measured to the closest 1mm by a portable stadiometer (Seca 123, Hamburg, Germany) and body weight was measured with the use of a digital scale (Seca 899), to the closest 0.1 kg. Waist circumference was measured with a measuring tape (Seca 210), where the students were asked to distribute their bodyweight equally on both legs, keep their hands alongside their body and to breathe normally throughout the measuring period. Participants were asked to be dressed in a t-shirt and pants, wearing no shoes, during all anthropometrical measures. An average of two waist measures, between the anterior iliac crest and the lower rib was then estimated. Lastly, BMI was calculated based on the student's weight and height measurements. All measurements were performed adhering to a standardized protocol.

Mental health difficulties

In this study, mental health difficulties were subjectively measured using the Strength and Difficulties Questionnaire. The questionnaire included a total of 25 questions, aimed to evaluate different aspects of mental health, categorized into five subscales (5 questions for each subcategory): Emotional symptoms, Conduct problems, Peer relationship problems, Hyperactivity and Prosocial behavior (Goodman, 1997).
As a three point likert scale, students were asked to tick off either of three answers that best represented their perceived state of mental health at that time, ranging from "Not

true", "Partly true" and "True". Each subscale was subsequently scored with a number of 0-10. Lastly, scores belonging to each subscale, with the exception of prosocial behaviours, was then summarized into a total mental health difficulty score, scaling from 0-40 where a higher score represent a high degree of symptoms of mental health problems. Students whom reported to have experienced mental health difficulties were asked to evaluate the degree in which this had impacted their daily living, by answering associated questions with either "not at all" (0 points), "only little" (0 points), "a good amount" (1 point) or "a lot" (2 points). This score was further summarized as an impact score.

Statistical methods

All participants whom reported valid data at baseline or follow-up by the SDQ questionnaire was included in the main analysis. Descriptive data are presented as mean and standard deviation (SD). Linear regression was applied to analyze the differences between the baseline measures obtained by students in the respective intervention arms. Normality and homogeneity of all sample data was assessed prior to the analysis. A linear mixed model with SDQ as the dependent variable was used for the main analysis, with a group-by-time interaction, and with individual mean values as the function of time. Since our randomized unit being schools, a "random effect" was added to the mixed model to accommodate the clustering of students within these units. Effect analysis is presented as adjusted mean for changes between interventional (M1 and M2) and control group (95% CI). Further, an exploratory analysis was conducted to test the effect moderation by gender and baseline levels of SDQ total score tertiles (sub-group). This analysis was performed using a linear mixed model including group (interventions and control) and SDQ total score subgroups (lower, middle and high tertile for SDQ total score baseline). Statistically significant interaction between genders was evident in several SDQ subscales ($p < 0.001$ for interaction), consequently all analyses were stratified by gender. All statistical analysis was performed using SPSS (IBM Corp. Released 2017. IBM SPSS Statistic for Windows, Version 25.0. Armonk, NY: IBM Corp), where the significance level was set to $p < 0.05$.

4.0 Results

4.1 Participants

Table 1 shows baseline characteristic of the adolescents stratified by gender and intervention groups. Both genders were equally distributed in each interventional models and control group. Boys in the M2 intervention group was on average 2,12 cm taller (P = 0.024), 4.2 kg heavier (P<0.001), had a 3,51 cm bigger waist circumference (P<0.001) and 0.94 points higher BMI score (P = 0.010), compared with the boys in the control group. No significant difference was seen between the boys in M1 or the girls in the two interventional groups, compared to the control group.

Table 1. Student's anthropometrical and demographical characteristic at baseline. Numbers are represented as mean and standard deviation (SD).

	Intervention (M1)		**Intervention (M2)**		**Control**	
	Girls	**Boys**	**Girls**	**Boys**	**Girls**	**Boys**
Students (n)	115	118	120	116	160	195
Age (yrs)	13,9 (0,3)	13,9 (0,27)	13,9 (0,3)	14,05 (0,3)	14 (0,3)	14 (0,3)
Height (cm)	162,8 (6,5)	166,6 (8,9)	164,1 (5,6)	169 (8,3)	164,1 (6,3)	166,8 (7,6)
Weight (kg)	52,8 (8,9)	53,4 (10,5)	55,1 (8,9)	57,4 (12,3)	54,3 (8,7)	52,9 (10,6)
BMI (kg/m^2)	19,92	19,23	20,46	20,09	20,16	19,01
Waist (cm)	65,2 (5,6)	66,8 (7,6)	66,6 (6,0)	69,9 (8,8)	66,3 (5,7)	66,5 (8,1)

Intervention M1= Physical active based learning; Intervention M2=Don't worry be happy; BMI=Body mass index

4.2 Intervention effects on self-reported SDQ categories of Mental health

Total symptoms of mental health problem score

Mental health problem scores at baseline and follow-up are presented in table 3 for girls and table 4 for boys.

The main analysis presented no significant intervention effect from neither intervention model on the total score of symptoms of mental health problems nor impact score, among girls (table 3) or boys (table 4) by the SDQ.

Symptoms of mental health problem subscales

When analyzing the different SDQ subcategories a significant intervention effect was seen in the "peer problems" category. The intervention effect was in favor of the boys in the M1 intervention model (mean difference in change -0.44, 95% CI -0,87 ; -0,01, p = 0.045) when compared with boys in the control group. In contrast, the intervention effect among girls was not in favor for the students in M2 (mean difference in change 0.55, 95% CI 0,12 ; 0,98, p = 0.012) when compared with girls in the control group.

Additionally, a significant positive effect was observed by the subcategory of prosocial behavior amongst boys in M1 intervention model when compared with their control school counterparts (mean difference in change 0.55, 95% CI 0,07 ; 1,02, p = 0.024). No significant effect was seen in the subcategory of hyperactivity, conduct or emotional symptoms in neither intervention groups, compared with control, in both genders.

Table 3. Adjusted mean values (95% CI) of SDQ measures at test 1 and test 2, and group differences of changes in SDQ total symptoms of mental health problem score, SDQ impact score, emotional symptoms, conduct problems, hyperactivity, relationship problems with peers, and prosocial behaviors for girls.

	Intervention (M1)		Intervention (M2)		Control		Group differences	
	Test 1	Test 2	Test 1	Test 2	Test 1	Test 2	M1 -C	M2 - C
Total Score	11 (9,59 ; 12,47)	11,36 (9.9 ; 12,8)	11,5 (10,1 ; 12,8)	12.1 (10,7; 13.56)	11,18 (9.9 ; 12,39)	11,65(10,4 ; 12,9)	-0,14 (-1,31 ; 1,02)	0,18 (-1,01 ; 1,37)
Impact score	0,02 (-0,08 ; 0,1)	0,16 (0,05 ; 0,26)	0,11 (0,01 ; 0,20)	0,18 (0,07 ; 0,29)	0,15 (0,06 ; 0,23)	0,14 (0,05 ; 0,23)	0,14 (-0,01 ; 0,29)	0,079 (-0,07 ; 0,23)
Emotion	3,60 (3,15 ; 4,06)	3,95 (3,49 ; 4,40)	4,27 (3,84 ; 4,70)	4,26 (3,78 ; 4,73)	4,10 (3,73 ; 4,47)	4,36 (3,97 ; 4,75)	0.09 (-0,45 ; 0,63)	-0,26 (-0,81 ; 0,28)
Conduct	1,45 (1,09; 1,82)	1,36 (1,00 ; 1,73)	1,25 (0,91 ; 1,59)	1,36 (1,00 ; 1,73)	1,45 (1,15 ; 1,75)	1,43 (1,12 ; 1,74)	-0,07 (-0,42 ; 0,28)	0,12 (-0,22 ; 0,48)
Hyperactivity	4,12 (3,67; 4,57)	4,27 (3,82 ; 4,71)	4,10 (3,68 ; 4,53)	4,31 (3,85 ; 4,77)	3,73 (3,37 ; 4,10)	4,17 (3,79 ; 4,55)	-0,29 (-0,79 ; 0,21)	-0,22 (-0,73 ; 0,27)
Peer	1,79 (1,32 ; 2,25)	1,74 (1,28 ; 2,2)	1,74 (1,32 ; 2,17)	2,06 (1,60 ; 2,52)	1,89 (1,51 ; 2,28)	1,65 (1,25 ; 2,05)	0.19 (-0,22 ; 0,62)	0,55 (0,12 ; 0,98)*
Prosocial behavior	7,86 (7,51 ; 8,22)	7,97 (7,61 ; 8,32)	8,4 (8,07 ; 8,73)	8,3 (7,94 ; 8,66)	8,2 (7,9 ; 8,49)	8,03 (7,72 ; 8,33)	0,27 (-12 ; 0,66)	0,06 (-0,33 ; 0,47)

M1=Active learning intervention; M2=Don’t worry – Be happy intervention.

* = $p < 0.05$

Table 4. Adjusted mean values (95% CI) of SDQ measures at test 1 and test 2, and group differences of changes SDQ total symptoms of mental health problem score, SDQ impact score, emotional symptoms, conduct problems, hyperactivity, relationship problems with peers and prosocial behaviors score for boys.

	Intervention (M1)		Intervention (M2)		Control		Group differences	
	Test 1	Test 2	Test 1	Test 2	Test 1	Test 2	M1 - C	M2 - C
Total Score	10.98 (9.83 ; 12.14)	10,8 (9,7 ; 11,9)	9,65 (8,54 ; 10,76)	10,19 (8,9 ; 11,4)	10,9 (10,02 ; 11,81)	11,6 (10,7 ; 12,5)	-0,87 (-2,14 ; 0,39)	-0,19 (-1,55 ; 1,17)
Impact score	0,03 (-0,04 ; 0,11)	0,09 (0,01 ; 0,16)	0,05 (-0,02 ; 0,12)	0,09 (0,13 ; 0,18)	0,08 (0,2 ; 0,13)	0,05 (-0,01 ; 0,1)	-0,08 (-0,03 ; 0,21)	0,07 (-0,05 ; 0,20)
Emotion	2,62 (2,14 ; 3,10)	2,60 (2,13 ; 3,07)	2,27 (1,82 ; 2,72)	2,24 (1,74 ; 2,73)	2,54 (2,15 ; 2,92)	2,82 (2,42 ; 3,21)	-0,30 (-0,81 ; 0,20)	-0,31 (-0,86 ; 0,23)
Conduct	2,08 (1,73 ; 2,43)	1,94 (1,60 ; 2,27)	1,65 (1,30 ; 2,00)	1,93 (1,55 ; 2,32)	2,11 (1,86 ; 2,36)	2,42 (2,15 ; 2,69)	-0,46 (-0,93 ; 0,008)	-0,02 (-0,53 ; 0,47)
Hyperactivity	4,05 (3,62 ; 4,48)	4,46 (4,05 ; 4,88)	3,91 (3,48 ; 4,34)	4,03 (3,55 ; 4,50)	4,22 (3,90 ; 4,53)	4,26 (3,92 ; 4,59)	0,37 (-0,16 ; 0,91)	0,07 (-0,50 ; 0,65)
Peer	2,22 (1,87 ; 2,57)	1,83 (1,49 ; 2,17)	1,81 (1,46 ; 2,16)	1,98 (1,60 ; 2,37)	2,05 (1,79 ; 2,31)	2,11 (1,83 ; 2,38)	-0,44 (-0,87; -0,01)*	0,12 (-0,34 ;0,58)
Prosocial behavior	7,26 (6,87 ; 7,65)	7,69 (7,32 ; 8,07)	7,58 (7,20 ; 7,97)	7,50 (7,07 ; 7,92)	7,56 (7,27 ; 7,86)	7,44 (7,13 ; 7,75)	0,55 (0,07 ; 1,02)*	0,03 (-0,48 ; 0,54)

M1=Active learning intervention; M2=Don't worry – Be happy intervention.
* = $p < 0.05$

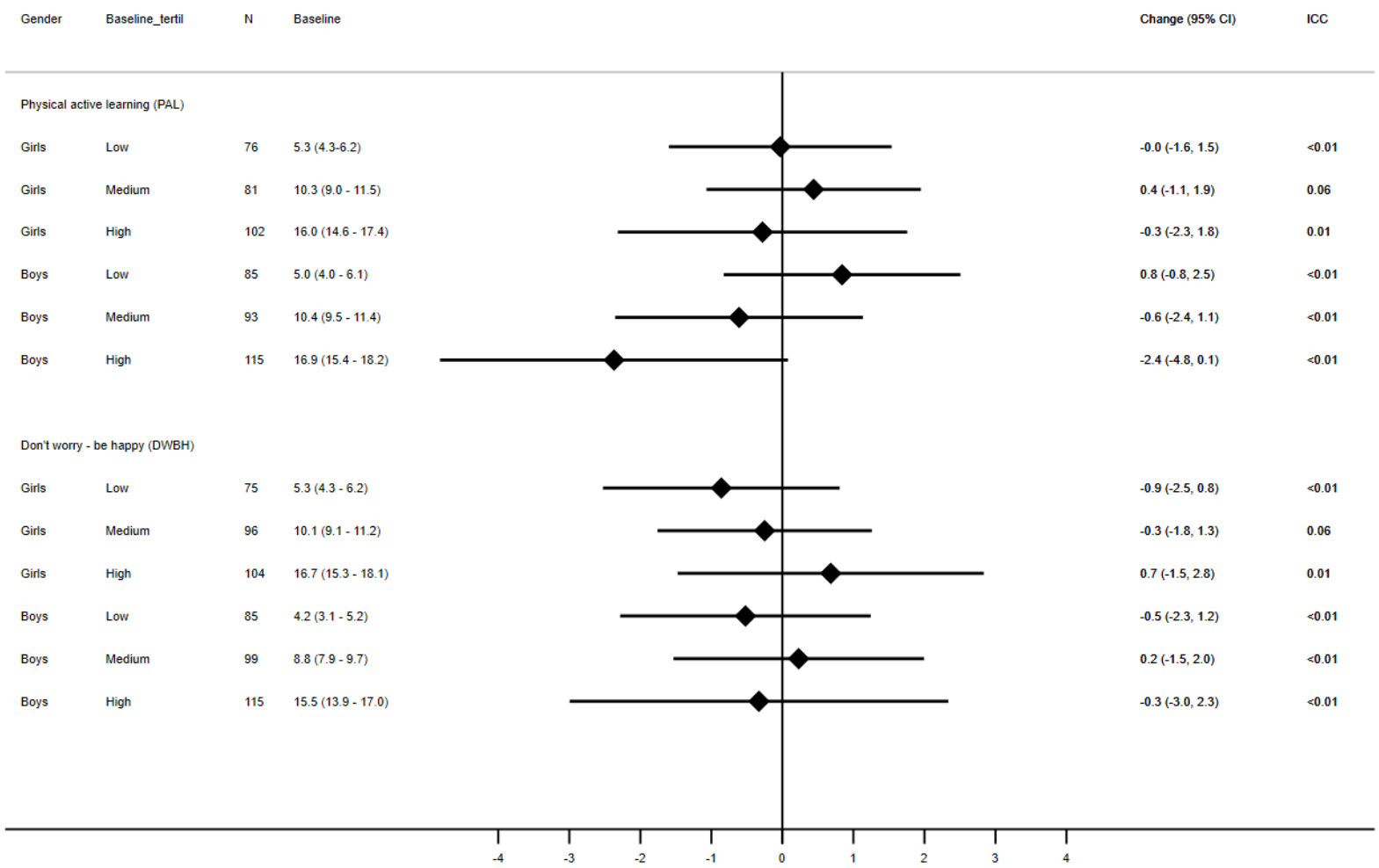
4.3 Degree of mental difficulties

Total score of symptoms of mental health problems (95% CI) was categorized in low, medium or high at baseline for all participants (figure 1). We found no significant differences in any of the three mental health difficulty subgroups at baseline, for neither interventional groups among either boys or girls when compared to the controls.

4.4 Total adherence rate

The total adherence rate amongst intervention schools were 83% for the M1 intervention and 81% for the M2 intervention. Based on the adherence rate, schools belonging to the M1 intervention would have completed a total of 100 minutes of intervention based activities on a weekly basis, whereas schools allocated to the M2 intervention completed a total of 97 minutes.

Graph 1. Intervention effect (95% CI) of M1 and M2 compared to control in degrees of total mental health problem for boys and girls.



M1=Active learning intervention; M2=Don't worry-Be happy intervention. Low=Low score of mental health difficulties; Med=Medium score of mental health difficulties; High=high score of mental health difficulties.

5.0 Discussion

The main purpose of this thesis was to investigate the effect of a 9-month long school-based PA intervention on symptoms of mental health problems amongst 14-year-olds adolescents in Akershus county, Norway. As a secondary purpose, the thesis examined whether the intervention effect varied in respect to the adolescents total SDQ score at baseline.

5.1 Main findings

We found no significant effects on total symptoms of mental health problem score by SDQ among adolescents in either intervention group when compared to their control school counterparts. This finding was similar in both genders. Further, we observed no significant intervention effect when we stratified the sample in subgroups of total SDQ score at baseline among either girls or boys.

Our findings correspond with the results of previous literature examining the effect of school-based PA interventions on total SDQ score among children and adolescents (Eather et al., 2015; Bunketorp et al., 2016). In these interventions, PA was integrated into the curriculum in various ways resulting in no intervention effect on total SDQ score. This includes an eight-week intervention, where PE classes (120 minutes) were substituted by high intensity strength-based exercise sessions, amongst 15-year-old adolescents (n = 51 intervention, n = 45 control) (Eather et al 2015). In the study by Bunketorp and colleagues (2016), two additional PE classes of 30-45 minutes was integrated on a weekly basis, for one school semester, focusing on playful activities and sport, amongst children enrolled in 0-6th grade (n = 182 intervention, n = 167 control).

Several observation studies have assessed the associative benefits linked to positive PA behaviors, both self-reported and objectively measured, on total SDQ score also. Amongst these studies, little significant associations have been found between a greater volume of PA and a lower total SDQ score (Vedøy et al. 2020; Sagatun et al. 2007; Bell et al. 2019; Wiles et al. 2008; Hallal et al. 2015). This includes the results from the longitudinal studies by Sagatun et al (2007),

as well as Bell & colleagues (2019) whom both assessed PA by accelerometry, and its relation to adolescents total SDQ score, three years later. Correspondingly, Wiles et al (2008) and Hallal et al (2015), found no association between adolescents whom self-reported to adhere to the recommended amount of 300 min/week of PA, and a lower total SDQ score, one and three years later respectively. Similarly, a null finding was also reported amongst 599 adolescents in the cross-sectional study by Vedøy and colleagues (2020).

One exception was the findings by Kuiper and colleagues (2018), whom reported a significant lower total SDQ score in active, compared to inactive self-reported PA levels amongst 96 617 adolescents (Kuiper and colleagues, 2018). Although mostly presenting similar null findings in regards to PA and its association with total SDQ score in adolescents, comparing our results with observational studies should be done with caution, as our study investigated the effect of implementing additional PA through interventions on mental health problems, and not the association between the two variables.

Compared to our study, significant effects by interventions have been reported in studies who have investigated the effect of various exercise and PA intervention modalities, on more specific mental health problem outcomes, with a majority of them assessing depression and anxiety symptoms solely (Biddle et al. 2019; Brown et al. 2013). Amongst these, the study sample is often seen consisting of adolescents whom are more greatly disposed to factors related to poor mental health, subsequently leading to greater mental health related problems at baseline. This includes adolescents with poor socioeconomic backgrounds, male youth offenders housed in institutions, and adolescents whom classified as being overweight (Brown et al., 2013; Bonhauser et al., 2005).

5.2 Physical activity dose

In respect to dosage, the two ScIM intervention arms intended to implement a total of 120 minutes of additional curricular based PA, each week. The total adherence from the participating schools allocated to the two intervention arms were 83% and 81 % for the M1 and M2 interventions, respectively.

In reality, this means that our adolescents completed on average a total of 100 min and 97 min of intervention based activities, on a weekly basis. Consequently, despite facilitating for additional PA opportunities, the dosage provided in our study may be viewed as a minor contributor to the overall PA volume on a weekly basis. To put it in to perspective, adolescents are recommended to accumulate 300 minutes of MVPA, per week, in order to support further health benefits (Parrish et al., 2020).

However, a recommended amount of PA in order to support mental health in adolescents has not been discovered yet. In the review of reviews by Biddle and colleagues (2019), neither frequency nor duration seemed to moderate the effect of PA and exercise on depression, nor anxiety scores across several meta analytic studies (Biddle et al 2019). Similarly, it is not clear whether the amount of PA play a significant role in lowering general symptoms of mental health problems, as measured by the SDQ. Most studies, mainly limited to observational studies, reports little significant relationship between performing a greater PA amount, and a lower total symptoms of mental health problems by SDQ (Kuiper, Broer, & Van, 2018; Vedøy et a,. 2020; Sagatun et al., 2007; Bell et al., 2019; Wiles et al., 2008; Hallal et al., 2015). However, its is important to emphasize that previous observational studies are hampered by the use of self-reported measures of PA. In these cases the adolescents are asked to recall their past PA behaviors in order to answer one or more questions related to the amount of activities performed on a weekly basis. In turn, recalling past behaviors are known to be a methodological weakness in most studies, as it yields a greater risk of recall bias when implemented amongst adolescents (Kavanaugh, Moore, Hibbett & Kaczynski, 2015). Nevertheless, some studies have implemented objective measures of PA, where a similar lack of association have been reported despite implementing measurements of greater quality (Bell et al., 2019; Piggin, Mansfield & Weed, 2017). Conclusively, based on the lack of evidence, it is difficult to argue whether any changes to our intervention dose would have produced any significant effect. As emphasized by the literature, more studies are needed in order to investigate the moderating effect of PA on mental health problems overall (Biddle et al., 2019).

5.3 Intervention duration

The duration of our interventions lasted for a total of 9 months and can be considered as being of considerable length, when compared to previous PA intervention studies with a similar measurement outcome (Eather et al; Bunketorp et al., 2016).
However, despite having a greater intervention length, our study did not show any different effect, compared to the eight week-long intervention by Eather and colleagues (2015), and the one semester long intervention by Bunketorp et al (2016), in total SDQ score. Again, little research has been conducted on whether a longer intervention duration may yield a greater effect on genereral symptoms of mental health problems. Therefore, it may be difficult to argue whether our intervention would have benefited from a longer implementation overall. Amongst studies investigating depression specific outcomes, Brown and colleagues (2013) reported in their meta analysis, to find a greater effect amongst shorter intervention durations (Brown et al., 2013; Biddle et al., 2019). To elaborate, the effect of PA and exercise on the outcome of depression symptoms, proved to be stronger amongst the few intervention studies that lasted for less than a three-month period, compared to longer interventions (Brown et al. 2013). However, in this review there was only two studies which lasted for less than three months, whereas seven studies lasted for longer (Brown et al., 2013). Therefore limited assumptions can be made in regards to whether the greater effect was indeed caused by having a shorter duration, or if it was caused by other factors. Overall, there is a clear lack of evidence in which supports the intervention duration to mediate any effects on mental health problems, which emphasize the need for further investigation. As a consequence of this, little assumptions can be made in regards to whether our intervention effect would have been greater based on the intervention duration itself.

5.4 Baseline levels

At baseline, our study sample reported a low mean score of subjectively measured mental health problems by the SDQ. This suggests that our adolescents experienced little mental health problems prior to participating in the two intervention arms. Because of this, we may argue that the potential effect by the two interventions may have been limited, as our adolescent's mental health may not have been in need for interventional treatment from the beginning.

A similar case was also reported by Eather and colleagues (2015), whom argued that the lower levels of ill-being amongst the study sample may have enhanced the risk of a ceiling effect, and therefore limiting the effect on mental health outcomes (Eather et al., 2015). This corresponds well with literature, as some studies have reported to find a larger intervention effect amongst adolescents with greater depression symptoms, when compared with a healthy population (Brown et al., 2015; Biddle et al., 2019).

Furthermore, this may also explain the lack of difference seen in the effect amongst the three SDQ tertiles at baseline. The upper limit by participants whom belonged to the subgroup of "high" symptoms of mental health problems, for both M1 and M2, scored an average of 16 points on the SDQ. In respect to the SDQ, a total score above 20 points signifies an abnormal amount of mental health problems (Goodman, 1997), whereas this threshold was not reached by our adolescents. Consequently, we may argue that the lack of intervention effect on total symptoms of mental health problems in our study, may be due to our participant's mental health not being "ill" enough (Eather et al. 2015).

5.5 Methodological factors

The literature has emphasized the importance of further conducting intervention-based studies with high methodological quality to assess the effect of PA on mental health (Biddle et al. 2019; Brown et al., 2013). There are numerous methodological factors that may be viewed as strength and weaknesses following its implementation. Also, highlighting these methodological factors are of great importance as they may have influenced the results by this study.

5.5.1 Study design

Our study was conducted as a randomized control trial, which is recognized as the gold standard when evaluating cause and effect amongst more controlled settings (Grimshaw, Campbell, Eccles & Steen 2000; Spieth et al., 2016). There have been few studies investigating the effect of a school-based PA intervention on general symptoms of mental health problems measured by SDQ, with an experimental RCT design, in adolescents (Eather et al., 2015).

As a strength, the study design allowed each participating school to be randomly allocated to either the two intervention groups or control. By doing so, this enabled the likelihood of equal dispersion of confounding factors, in which could potentially mediate or moderate the effect between intervention and control groups. Conclusively, the likelihood of a biased results following the intervention was minimized, therefore improving the internal validity of our study overall (Spieth et al., 2016).

However, our study design did not include any blinding of neither participants, nor researchers. This could potentially have weakened the internal validity and the results of the ScIM study, as a full awareness of the intervention and its function could have affected behaviors amongst participants involved (Schulz & Grimes, 2002).

5.5.2 Study sample

The study in which this thesis is based on included a considerable large sample of participants when compared to previous intervention studies whom included a similar measurement outcome (Eather et al., 2015; Bunketorp et al., 2016). A total of 1149 students was invited from 11 different schools, where a total of 886 adolescence provided a signed concent, equivalating to a participation rate of 77% from our overall study sample. Consequently, the high percentage of participation from each school and the overall sample size, may be seen as a methodological strength, as it can decrease the likelihood of random error to occur (Stephenson & Babiker, 2000). For the original ScIM study, schools were invited based on a random pick from a list of schools surrounding the geographical areas of the four test centers. In addition, participating schools were invited based on several pre-determined inclusion and exclusion criteria. As a consequence of this, the study was more likely to recruit a study sample that better represent a more general Norwegian adolescent population, therefore enhancing the studies generalizability (Spieth et al., 2016).

In respect to the internal validity, it is important to emphasize that our study sample did not meet the size requirement stated by the power analysis performed by the ScIM study. This analysis tested the amount of participants needed in order to detect changes in PA levels.

Based on this, we may assume that the size of our sample did not have a sufficient power to confidently detect changes in total SDQ score as well. Consequently, a lack of power could be considered as a weakness to the internal validity, and the results by this thesis (Schulz & Grimes, 2005). Nevertheless, a power analysis was not conducted on the SDQ variable for our sample, therefore, assumptions can only be made.

Also, values concerning SDQ variables from either baseline or at follow up was missing amongst some participants. To best handle the the dataset, a mixed model was administered to analyse the differences between the two intervention arms up against the control. By doing so our thesis was able to include all reported SDQ values in its analysis, by regressing its value by the variables of our two testpoints and the intervention arms. As a methodological strength, analyzing all values regardless of dropout could minimize the likelyhood of a biased result in our study (Xi, Pennel, Andridge & Paskett, 2018). Alternatively, intention to treat analysis is often seen used by RCT studies (Gupta, 2011). This analysis is commonly used to detect changes amongst groups, based on intervention allocations, regardless of withdrawals or protocol deviations by participants (Gupta, 2011). However, the intention-to-treat analysis have been known to lack the ability to handle missing values sufficiently, which in our case, justifies the use of a mixed model analysis to cater for our missing data (Gupta, 2011).

5.5.3 ScIM intervention models

Physical activity-based learning (M1)

One of the functions of the M1 intervention, besides enhancing overall PA levels, was to deliver more high intensity activities to the adolescents on a weekly basis. Adolescents would perform sporting activities of interest, that would lead to exhaustion. Subsequently, we may suggest that the M1 intervention aimed to facilitate physical health and its associative responses. How PA may affect mental health in adolescents are less known, and the need for further scientific development is emphasized by the literature (Doré et al., 2020; Lubans et al., 2016). However, in relation to previous mechanisms hypothesized to explain this effect, the M1 intervention may be

facilitating for neurobiological responses. With this, PA is hypothesized to impose positive acute and chronic effects on mental health and cognitive functioning (Lubans et al., 2016). In the light of the M1 intervention, the neurobiological mechanism suggests that the adolescents may experience physiological benefits from the increased PA volume and intensity. More specifically, a release of noradrenaline, serotonin and dopamine may occur following the bouts of activities, which have been hypothesized to impose a positive effect, and to relieve mental health problems acutely (Lubans et al., 2016). Likewise, a release of endorphins following the activities, have shown to provide the feeling of euphoria (Lubans et al., 2016).
However, little is known whether these acute responses yield any long-term effect on mental health. Therefore, it is unsure whether the interventions could have produced any lasting effects on the adolescent's mental health, which in our case may not have occurred.

Furthermore, the M1 intervention implemented an additional PE class, mainly focusing on sporting activities. Participating in group-based sporting activities have shown to enhance social connectiveness with peers, and with a possibility to facilitate for self-efficacy and confidence amongst adolescents (Doré et al., 2020; Lubans et al., 2016). In turn, these factors may further moderate positive effects on mental health outcomes (Doré et al., 2020; Lubans et al., 2016). In fact, the M1 intervention proved to have a significant positive effect on the subcategories of prosocial behaviors and a lower score of peer related problems amongst boys, compared to control ($P = 0.024$ and $P = 0.045$). This may suggest that the boys in M1 did experience a greater social connectiveness through the additional sporting activities, which potentially could lead to lower levels of mental health problems. However, the positive effect on the two SDQ subcategories did not seem to be great enough in order to influence the total SDQ score.

Don't worry-Be happy (M2)

The second intervention DWBH (M2) aimed to facilitate additional activities, in a way that would allow the adolescents to participate based on their own interests and competency. The purpose of the intervention was to support new friendships and to act as a social platform, whilst giving the adolescents the ability to freely choose activities. Consequently, the aim of the intervention may be supporting the basic needs theory, including autonomy, self-efficacy and

relatedness with peers (Lubans et al., 2016; Doré et al., 2020). Fulfilling these three basic needs are known to support psychological well-being and motivation amongst adolescents. Similarly, the literature has suggested that social relationships is a key factor in which could potentially moderate mental health (Doré et al., 2020). However, in relation to school-settings and curricular based activities, involvement is often seen as being mandatory. Due to this, we may argue as to whether the adolescents participated based on personal belief and own choice. As an alternative, the intervention may not have been efficient in supporting autonomy amongst the adolescents, and in turn limited the involvement and enjoyment of the activities. This could have potentially impaired the effect of the M2 intervention due to the lack of motivation by the adolescents.

Likewise, social relationships amongst adolescents is known to tested during this period of physical and biological changes (Herpertz-Dahlmann, 2013). Therefore, supporting relatedness to peer's and developing new friendships may be challenging during adolescence. Interestingly, the intervention did prove to have a negative effect on peer problems amongst girls in the M2 (DWBH) group, in which they reported a higher score, compared with the control ($P = 0.012$). This may suggest that the unsupervised and unstructured nature of the M2 intervention, could have caused additional strains in the social relationships amongst the girls. Also it is unknown whether the intervention facilitated new relationships between adolescents, or if they grouped up with a familiar social surrounding, which would have weakened the purpose of facilitating for new social relationships in the M2 intervention. Furthermore, as the intervention left the adolescents mostly unsupervised, little is known whether they performed a sufficient amount of activities during this school-based hour. There may have been a possibility that the adolescents used this opportunity as a break, which in turn would have contradicted the overall purpose of the intervention itself.

5.5.4 SDQ measurement

Despite having been previously validated in numerous countries, studies have reported that the Norwegian translated version of SDQ may have some weaknesses, and that its validity can only be supported to some extent (Ritcher, Sagatun, Heyerdahl, Oppedal & Røysamb 2011). As a general mental health questionnaire, the SDQ aims to measure commonly known symptoms in

which are linked to mental health problems in adolescence (Deighton et al., 2014). As a consequence of this, the SDQ is known to be less sensitive in its measure, compared with more diagnostic specific questionnaires (Deighton et al., 2014). In fact, Brown and colleagues (2015) reported a greater intervention effect amongst studies whom focused on depression only as an outcome measure, compared to more general mental health outcomes (Brown et al 2015). This could suggest that using a more specific measure in which targets one diagnosis only, provides a more sensitive measure. As a consequence, a more specified measure could be more proficient in detecting changes over time, compared to general symptoms questionnaires, such as the SDQ (Deighton et al., 2014).

Amongst adolescents, mental health problems have been known to be poorly reported, partly due to its stigmatic nature (Pascoe et al., 2020). Therefore, we may argue to what degree the participants reported a true score of symptoms of mental health problems by the SDQ. If the adolescents depicted a false representation of their mental health, this could have impaired the outcome effect of our study, as a biased measure would have been reported. Also, in the ScIM study, participants overall mental health was measured through a self-reported questionnaire. The overall questionnaire included other questions related to mental health in addition to the SDQ, in which the students used on average 30 min to complete. In turn, the overall length of the questionnaire can be seen as a weakness, as this may cause the adolescents to lose interest as well as focus, which in turn may impact the validity of the SDQ scoring in a negative manner.

6.0 Conclusion

This study has shown that a 9 month long, school-based PA intervention with two different intervention arms, had little significant effect on total symptoms of mental health problem scores in Norwegian 14 year old adolescents. Additionally, the intervention effect did not vary significantly in respect to the adolescent's total symptoms of mental health score measured at baseline. A significant effect was however seen in lowering relationship with peer problems, and increasing pro-social behaviors amongst boys, from the increased PE and PA sessions by the M1 intervention. Conversely, a significant negative effect was seen amongst the girls, where relationship with peer problems increased following two additional PA sessions on a weekly basis, aimed to enhance social interaction and with minimal involvement by teachers.

7.0 References

Bell, S. L., Audrey, S., Gunnell, D., Cooper, A., & Campbell, R. (2019). *The relationship between physical activity, mental wellbeing and symptoms of mental health disorder in adolescents: a cohort study*. International Journal of Behavioral Nutrition and Physical Activity, 16(1): 138.

Biddle, S. J., & Asare, M. (2011). *Physical activity and mental health in children and adolescents: a review of reviews*. British journal of sports medicine, 45(11): 886-895.

Biddle, S. J., Ciaccioni, S., Thomas, G., & Vergeer, I. (2019). *Physical activity and mental health in children and adolescents: An updated review of reviews and an analysis of causality*. Psychology of Sport and Exercise, 42: 146-155.

Bonhauser, M., Fernandez, G., Püschel, K., Yañez, F., Montero, J., Thompson, B., & Coronado G. (2005). *Improving physical fitness and emotional well-being in adolescents of low socioeconomic status in Chile: results of a school-based controlled trial*. Health Promotion International, 20(2): 113-122.

Bouten, C. V., Koekkoek, K. T., Verduin, M., Kodde, R., & Janssen, J. D. (1997). *A triaxial accelerometer and portable data processing unit for the assessment of daily physical activity*. IEEE transactions on biomedical engineering, 44(3): 136-147.

Brage, S., Brage, N., Ekelund, U., Luan, J. A., Franks, P. W., Froberg, K., & Wareham, N. J. (2006). *Effect of combined movement and heart rate monitor placement on physical activity estimates during treadmill locomotion and free-living*. European journal of applied physiology, 96(5): 517-524.

Brown, H. E., Pearson, N., Braithwaite, R. E., Brown, W. J., & Biddle, S. (2013). *Physical activity interventions and depression in children and adolescents*. Journal of Sports medicine, 43(3): 195-206.

Bunketorp, K. L., Malmgren, H., Olsson, E., Lindén, T., & Nilsson, M. (2015). *Effects of a curricular physical activity intervention on children's school performance, wellness, and brain development*. Journal of School Health, 85(10): 704-713.

Caspersen, C. J., Powell, K. E., Christenson, G. M. (1985). *Physical activity, exercise, and physical fitness: definitions and distinctions for health-related research*. Public Health Rep. 1985;100(2): 126–-31.

Das, J. K., Salam, R. A., Lassi, Z. S., Khan, M. N., Mahmood, W., Patel, V., & Bhutta, Z. A. (2016). *Interventions for adolescent mental health: an overview of systematic reviews*. Journal of Adolescent Health, 59(4): S49-S60.

Deighton, J., Croudace, T., Fonagy, P., Brown, J., Patalay, P., & Wolpert, M. (2014). *Measuring mental health and wellbeing outcomes for children and adolescents to inform practice and policy: a review of child self-report measures*. Child and adolescent psychiatry and mental health, 8(1): 14.

Derdikman-Eiron, R. (2012). *Symptoms of Anxiety and Depression and Psychosocial Function in Males and Females from Adolescence to Adulthood: Longitudinal Findings from the Nord-Trøndelag Health Study.*

Doré, I., Sylvester, B., Sabiston, C., Sylvestre, M. P., O'Loughlin, J., Brunet, J., & Bélanger, M. (2020). *Mechanisms underpinning the association between physical activity and mental health in adolescence: a 6-year study*. International Journal of Behavioral Nutrition and Physical Activity, 17(1): 9.

Duinhof, E. L., Lek, K. M., de Looze, M. E., Cosma, A., Mazur, J., Gobina, I., ... & Stevens, G. W. J. M. (2020). *Revising the self-report strengths and difficulties questionnaire for cross-country comparisons of adolescent mental health problems: the SDQ-R*. Epidemiology and psychiatric sciences, 29.

Eather, N., Morgan, P., & Lubans, D. (2016). *Effects of exercise on mental health outcomes in adolescents: findings from the CrossFit™ teens randomized controlled trial*. Journal of Psychology of Sport Exercsie, 26: 14-23.

Essau, C. A., Olaya, B., Anastassiou-Hadjicharalambous, X., Pauli, G., Gilvarry, C., Bray, D., ... & Ollendick, T. H. (2012). *Psychometric properties of the Strength and Difficulties Questionnaire from five European countries*. International Journal of Methods in Psychiatric Research, 21(3): 232-245.

Fairclough, S., & Lubans, D. (Eds.). (2020). *The Routledge Handbook of Youth Physical Activity*. Routledge.

Goodman, R. (2001). *Psychometric properties of the strengths and difficulties questionnaire.* Journal of the American Academy of Child Adolescent Psychiatry, 40 (11).

Goodman, R. (1997). *The strengths and difficulties questionnaire: A research note*. Journal of Child Psychology and Psychiatry and Allied Disciplines. 1997;38(5): 581-586

Goodman, R. & Scott, P. (1999). *Comparing the Strengths and Difficulties Questionnaire and the Child Behavior Checklist: is small beautiful?* Journal of abnormal child psychology, 27(1): 17-24.

Green, K., & Smith, A. (Eds.). (2016). *Routledge handbook of youth sport*. Routledge.

Grimshaw, J., Campbell, M., Eccles, M., & Steen, N. (2000). *Experimental and quasi-experimental designs for evaluating guideline implementation strategies*. Family practice, 17, S11-S16.

Guo, C., Tomson, G., Keller, C., & Söderqvist, F. (2018). *Prevalence and correlates of positive mental health in Chinese adolescents*. BMC Public Health, 18(1): 263.

Gupta, S. K. (2011). *Intention-to-treat concept: a review*. Perspectives in clinical research, 2(3), 109.

Hallal, P. C., Martínez-Mesa, J., Coll, C. V. N., Mielke, G. I., Mendes, M. A., Piexoto, M. B., . . . Gonçalves, H. (2015). *Physical Activity at 11 Years of Age and Incidence of Mental Health Problems in Adolescence: Prospective Study*. Journal of Physical Activity Health, 12(4): 535-539.

Hawes, D. J., & Dadds, M. R. (2004). *Australian data and psychometric properties of the Strengths and Difficulties Questionnaire*. New Zealand journal of Psychiatry, 38(8): 644-651.

Herpertz-Dahlmann, B., Bühren, K., & Remschmidt, H. (2013). *Growing up is hard: mental disorders in adolescence*. Deutsches Ärzteblatt International, 110(25): 432.

Hills, A. P., Dengel, D. R., & Lubans, D. R. (2015). *Supporting public health priorities: recommendations for physical education and physical activity promotion in schools*. Progress in cardiovascular diseases, 57(4): 368-374.

Kavanaugh, K., Moore, J. B., Hibbett, L. J., & Kaczynski, A. T. (2015). *Correlates of subjectively and objectively measured physical activity in young adolescents*. Journal of Sport and Health Science, 4(3): 222-227.

Keyes, C. L., Shmotkin, D., & Ryff, C. D. (2002). *Optimizing well-being: the empirical encounter of two traditions*. Journal of personality and social psychology, 82(6): 1007.

Klasen, H., Woerner, W., Wolke, D., Meyer, R., Overmeyer, S., Kaschnitz, W., ... & Goodman, R. (2000). *Comparing the German versions of the strengths and difficulties questionnaire (SDQ-Deu) and the child behavior checklist*. European child & adolescent psychiatry, 9(4): 271-276.

Koskelainen, M., Sourander, A., & Vauras, M. (2001). *Self-reported strengths and difficulties in a community sample of Finnish adolescents*. Jouranl of European Child Adolescent Psychiatry 10(3): 180-185.

Kuiper, J., Broer, J., & van der Wouden, J. C. (2018). *Association between physical exercise and psychosocial problems in 96 617 Dutch adolescents in secondary education: A cross-sectional study*. European Journal of Public Health, 28(3): 468-473.

Liu, M., Wu, L., & Ming, Q. (2015). *How does physical activity intervention improve self-esteem and self-concept in children and adolescents? Evidence from a meta-analysis*. PloS one, 10(8).

Lubans, D., Richards, J., Hillman, C., Faulkner, G., Beauchamp, M., Nilsson, M., ... & Biddle, S. (2016). *Physical activity for cognitive and mental health in youth: a systematic review of mechanisms.* Pediatrics, 138(3).

Parrish, A. M., Tremblay, M. S., Carson, S., Veldman, S. L., Cliff, D., Vella, S., ... & Aubert, S. (2020). *Comparing and assessing physical activity guidelines for children and adolescents: a systematic literature review and analysis*. International Journal of Behavioral Nutrition and Physical Activity, 17(1), 16.

Pascoe, M., Bailey, A. P., Craike, M., Carter, T., Patten, R., Stepto, N., & Parker, A. (2020). *Physical activity and exercise in youth mental health promotion: a scoping review*. BMJ open sport & exercise medicine, 6(1).

Piggin, J., Mansfield, L., & Weed, M. (Eds.). (2017). *Routledge handbook of physical activity policy and practice.* Routledge.

Richter, J., Sagatun, Å., Heyerdahl, S., Oppedal, B., & Røysamb, E. (2011). *The strengths and difficulties questionnaire (SDQ)–Self-Report. An analysis of its structure in a multiethnic urban adolescent sample*. Journal of Child Psychology and Psychiatry, 52(9), 1002-1011.

Ritchie, H. & Roser, M. (2018). *Mental Health. Our World in Data*. Retrieved May 19, 2020 from https://ourworldindata.org/mental-health.

Rodriguez-Ayllon, M., Cadenas-Sanchez, C., Estevez-Lopez, F., Munoz, N. E., Mora-Gonzalez, J., Migueles, J. H., ... & Catena, A. (2019). *Role of physical activity and sedentary behavior in the mental health of preschoolers, children and adolescents: a systematic review and meta-analysis*. Sports medicine, 1-28.

Rønning, J. A., Handegaard, B. H., Sourander, A., & Mørch, W. T. (2004). *The Strengths and Difficulties Self-Report Questionnaire as a screening instrument in Norwegian community samples*. European child & adolescent psychiatry, 13(2), 73-82.

Sagatun, A., Søgaard, A. J., Bjertness, E., Selmer, R., & Heyerdahl, S. (2007). *The association between weekly hours of physical activity and mental health: a three-year follow-up study of 15–16-year-old students in the city of Oslo, Norway*. BMC public health, 7(1), 155.

Schulz, K. F., & Grimes, D. A. (2005). *Sample size calculations in randomised trials: mandatory and mystical*. The Lancet, 365(9467), 1348-1353.

Schulz, K. F., & Grimes, D. A. (2002). *Blinding in randomised trials: hiding who got what*. The Lancet, 359(9307), 696-700.

Sirard, J. R., & Pate, R. R. (2001). *Physical activity assessment in children and adolescents*. Sports medicine, 31(6), 439-454.

Spieth, P. M., Kubasch, A. S., Penzlin, A. I., Illigens, B. M. W., Barlinn, K., & Siepmann, T. (2016). *Randomized controlled trials–a matter of design*. Neuropsychiatric disease and treatment, 12, 1341.

Steinhausen, H. C., & Metzke, C. W. (2001). *Risk, compensatory, vulnerability, and protective factors influencing mental health in adolescence*. Journal of Youth and Adolescence, 30(3), 259-280.

Stephenson, J. M., & Babiker, A. (2000). *Overview of study design in clinical epidemiology*. Sexually transmitted infections, 76(4), 244-247.

Theron, L., Liebenberg, L., & Malindi, M. (2014). *When schooling experiences are respectful of children's rights: A pathway to resilience*. School Psychology International, 35(3), 253-265.

Ueno, K. (2005). *The effects of friendship networks on adolescent depressive symptoms*. Social Science Research, 34(3), 484-510.

Vedøy, I. B., Anderssen, S. A., Tjomsland, H. E., Skulberg, K. R., & Thurston, M. (2020). *Physical activity, mental health and academic achievement: a cross-sectional study of Norwegian adolescents*. Mental Health and Physical Activity, 18, 100322.

Vos, T., Barber, R. M., Bell, B., Bertozzi-Villa, A., Biryukov, S., Bolliger, I., ... & Duan, L. (2015). *Global, regional, and national incidence, prevalence, and years lived with disability for 301 acute and chronic diseases and injuries in 188 countries, 1990–2013: a systematic analysis for the Global Burden of Disease Study 2013*. The Lancet, 386(9995), 743-800.

Wiles, N. J., Jones, G. T., Haase, A. M., Lawlor, D. A., Macfarlane, G. J., & Lewis, G. (2008). *Physical activity and emotional problems amongst adolescents*. Social psychiatry and psychiatric epidemiology, 43(10), 765.

Williams Jr, J. W., Noël, P. H., Cordes, J. A., Ramirez, G., & Pignone, M. (2002). *Is this patient clinically depressed?*. Jama, 287(9), 1160-1170.

World Health Organization (WHO). (2020a). *Mental disorders*. Retrieved June 23, 2020, from https://www.who.int/mental_health/management/en/

World Health Organisation (WHO). (2020b). *Adolescent mental health*. Retrieved June 23, 2020, from https://www.who.int/news-room/fact-sheets/detail/adolescent-mental-health

World Health Organization (WHO). (2004) P*romoting mental health: concepts, emerging evidence, practice*. Retrieved June 07, 2020, from http://www.who.int/mental_health/evidence/en/promoting_mhh.pdf.

Xi, W., Pennell, M. L., Andridge, R. R., & Paskett, E. D. (2018). *Comparison of intent-to-treat analysis strategies for pre-post studies with loss to follow-up*. Contemporary clinical trials communications, 11, 20-29.