Intervention and moderation of physical fitness in children with physical fitness deficits - Results of the SMaRTER study

An academic thesis submitted to the Faculty of Human Sciences of the University of Potsdam for the degree Doctor of Philosophy (PhD)

Fabian Arntz

2023-11-13

Online available under the following URL
https://faaa6.github.io/SMaRTER/index.html

Published online on the
Publication Server of the University of Potsdam:
https://doi.org/10.25932/publishup-62260
https://nbn-resolving.org/urn:nbn:de:kobv:517-opus4-622607

Table of contents

List of Figures

List of Tables

Abstract

Background: Physical fitness is a key aspect of children's ability to perform activities of daily living, engage in leisure activities, and is associated with important health characteristics. As such, it shows multi-directional associations with weight status as well as executive functions, and varies according to a variety of moderating factors, such as the child's gender, age, geographical location, and socioeconomic conditions and context. The assessment and monitoring of children's physical fitness has gained attention in recent decades, as has the question of how to promote physical fitness through the implementation of a variety of programs and interventions. However, these programs and interventions rarely focus on children with deficits in their physical fitness. Due to their deficits, these children are at the highest risk of suffering health impairments compared to their more average fit peers. In efforts to promote physical fitness, schools could offer promising and viable approaches to interventions, as they provide access to large youth populations while providing useful infrastructure. Evidence suggests that school-based physical fitness interventions, particularly those that include supplementary physical education, are useful for promoting and improving physical fitness in children with normal fitness. However, there is little evidence on whether these interventions have similar or even greater effects on children with deficits in their physical fitness. Furthermore, the question arises whether these measures help to sustainably improve the development/trajectories of physical fitness in these children.

The present thesis aims to elucidate the following four objectives: (1) to evaluate the effects of a 14 week intervention with 2 x 45 minutes per week additional remedial physical education on physical fitness and executive function in children with deficits in their physical fitness; (2) to assess moderating effects of body height and body mass on physical fitness components in children with physical fitness deficits; (3) to assess moderating effects of age and skeletal growth on physical fitness in children with physical fitness deficits; and (4) to analyse moderating effects of different physical fitness components on executive function in children with physical fitness deficits.

Methods: Using physical fitness data from the EMOTIKON study, 76 third graders with physical fitness deficits were identified in 11 schools in Brandenburg state that met the requirements for implementing a remedial physical education intervention (i.e., employing specially trained physical education teachers). The fitness intervention was implemented in a cross-over design and schools were randomly assigned to either an intervention-control or control-intervention group. The remedial physical education intervention consisted of a 14 week, 2 x 45 minutes per week remedial physical education curriculum supplemented by a physical exercise homework program. Assessments were conducted at the beginning and end of each intervention and control period, and further assessments were conducted at the beginning and end of each school year until the end of sixth grade. Physical fitness as the primary outcome was assessed using fitness tests implemented in the EMOTIKON study (i.e., lower body muscular strength (standing long jump), speed (20 m sprint), cardiorespiratory fitness (6 min run), agility (star run), upper body muscular strength (ball push test), and balance (one leg balance)). Executive functions as a secondary outcome were assessed using attention and psychomotor processing speed (digit symbol substitution test), mental flexibility and fine motor skills (trail making test), and inhibitory control (Simon task). Anthropometric measures such as body height, body mass, maturity offset, and body composition parameters, as well as socioeconomic information were recorded as potential moderators.

Results: (1) The evaluation of possible effects of the remedial physical education intervention on physical fitness and executive functions of children with deficits in their physical fitness did not reveal any detectable intervention-related improvements in physical fitness or executive functions. The implemented analysis strategies also showed moderating effects of body mass index (BMI) on performance in 6 min run, star run, and standing long jump, with children with a lower BMI performing better, moderating effects of proximity to Berlin on performance in the 6 min run and standing long jump, better performances being found in children living closer to Berlin, and overall gendered differences in executive function test performance, with boys performing better compared to girls. (2) Analysing moderating effects of body height and body mass on physical fitness performance, better overall physical fitness performance was found for taller children. For body mass, a negative effect was found on performance in the 6 min run (linear), standing long jump (linear), and 20 m sprint (quadratic), with better performance associated with lighter children, and a positive effect of body mass on performance in the ball push test, with heavier children performing better. In addition, the analysis revealed significant interactions between body height and body mass on performance in 6 min run and 20 m sprint, with higher body mass being associated with performance improvements in larger children, while higher body mass was associated with performance declines in smaller children. In addition, the analysis revealed overall age-related improvements in physical fitness and was able to show that children with better overall physical fitness also elicit greater age-related improvements. (3) In the analysis of moderating effects of age and maturity offset on physical fitness performances, two unrotated principal components of z-transformed age and maturity offset values were calculated (i.e., relative growth = (age + maturity offset)/2; growth delay = (age - maturity offset)) to avoid colinearity. Analysing these constructs revealed positive effects of relative growth on performances in star run, 20 m sprint, and standing long jump, with children of higher relative growth performing better. For growth delay, positive effects were found on performances in 6 min run and 20 m sprint, with children having larger growth delays showing better performances. Further, the model revealed gendered differences in 6 min run and 20 m sprint performances with girls performing better than boys. (4) Analysing the effects of physical fitness tests on executive function revealed a positive effect of star run and one leg balance performance and a negative effect of 6 min run performance on reaction speed in the Simon task. However, these effects were not detectable when individual differences were accounted for. Then these effects showed overall positive effects, with better performances being associated with faster reaction speeds. In addition, the analysis revealed a positive correlation between overall reaction speed and effects of the 6 min run, suggesting that children with greater effects of 6 min run had faster overall reaction speeds. Negative correlations were found between star run effects and age effects on Simon task reaction speed, meaning that children with larger star run effects had smaller age effects, and between 6 min run effects and star run effects on Simon task reaction speed, meaning that children with larger 6 min run effects tended to have smaller star run effects on Simon task reaction speed and vice versa.

Conclusions: (1) The lack of detectable intervention-related effects could have been caused by an insufficient intervention period, by the implementation of comprehensive and thus non-specific exercises, or by both. Accordingly, longer intervention periods and/or more specific exercises may have been more beneficial and could have led to detectable improvements in physical fitness and/or executive function. However, it remains unclear whether these interventions can benefit children with deficits in physical fitness, as it is possible that their deficits are not caused by a mere lack of exercise, but rather depend on the socioeconomic conditions of the children and their families and areas. Therefore, further research is needed to assess the moderation of physical fitness in children with physical fitness deficits and, in particular, the links between children's environment and their physical fitness trajectories. (2) Findings from this work suggest that using BMI as a composite of body height and body mass may not be able to capture the variation associated with these parameters and their interactions. In particular, because of their multidirectional associations, further research would help elucidate how BMI and its subcomponents influence physical fitness and how they vary between children with and without physical fitness deficits. (3) The assessment of growth-related changes indicated negative effects associated with the growth spurt approaching age of peak height velocity, and furthermore showed significant differences in these effects between children. Thus, these effects and possible interindividual differences should be considered in the assessment of the development of physical fitness in children. (4) Furthermore, this work has shown that the associations between physical fitness and executive functions vary between children and may be moderated by children's socioeconomic conditions and the structure of their daily activities. Further research is needed to explore these associations using approaches that account for individual variance.

Technical information

This thesis and all it's analysis were written and rendered using RStudio (RStudio Team, 2022) and Quarto (Allaire et al., 2022).

Part I

Introduction

1 Review of literature

1.1 Physical fitness

The term physical fitness has served as an umbrella term for a variety of different concepts and perspectives. In this thesis, physical fitness is understood as a multidimensional concept that includes: (1) the ability to perform daily tasks with vigor and alertness, without excessive fatigue, and with sufficient energy to enjoy leisure activities and cope with unforeseen emergencies (Caspersen et al., 1985; Corbin et al., 2000; Howley & Franks, 1986; R. Pate et al., 2012; R. R. Pate, 1983) and (2) is related to current and future health status (Corbin et al., 2000; Howley & Franks, 1986; Malina, 2004; R. Pate et al., 2012). Thus, physical fitness is an umbrella term referring to several components, such as cardiorespiratory fitness, muscular fitness, motor fitness, flexibility, and body composition, each with further subcomponents. Table 1.1 provides an overview, including definitions of different (sub)components.

Table 1. 1: Components of physical fitness and definitions

fitness components	definition
cardiorespiratory fitness	ability of circulatory and respiratory systems to supply energy during sustained physical activity and to eliminate fatigue products after energy supply (Corbin et al., 1978)
muscular fitness	
muscular endurance	(1) ability of a muscle or muscle group to exert force over many repetitions or successive loads (Corbin et al., 1978) or (2) sustaining voluntary contractions over a prolonged period of time (Ortega et al., 2014)
muscular strength	ability of a muscle or muscle group to exert/generate an external force (Corbin et al., 1978; Ortega et al., 2014)
muscular power	ability of a muscle or muscle group to perform a maximum dynamic contraction in a short time (Ortega et al., 2014)
motor fitness	
agility	(1) ability to change position of the whole body in space quickly and precisely (Corbin et al., 1978; Ortega et al., 2014) or (2) a combination of speed, balance, strength, and coordination (Ortega et al., 2008)
balance	ability to maintain balance while stationary or in motion (Corbin et al., 1978)
coordination	ability to use senses, such as vision and hearing, together with body parts to perform motor tasks smoothly and accurately (Corbin et al., 1978)
power	(1) ability, which refers to speed at which a person can perform work (Corbin et al., 1978), or (2) ability to exert force quickly (Howley & Franks, 1986)
reaction time	ability that refers to time elapsed between stimulation and onset of response (Corbin et al., 1978)

Table 1. 1: Components of physical fitness and definitions (continued)

fitness components	definition
speed	ability to perform movements in a short period of time (Corbin et al., 1978)
flexibility	ability of a muscle or group of muscles to move a joint freely through full range of motion (Corbin et al., 1978; Ortega et al., 2014)
body composition	health-related component referring to the relative amount of muscle, fat, bone, and other vital parts of the body (Corbin et al., 1978)

Cardiorespiratory fitness can also be referred to as cardiorespiratory endurance (Corbin et al., 1978), cardiorespiratory function (Howley & Franks, 1986), aerobic fitness (Léger et al., 1988; Tomkinson, 2007), cardiovascular fitness, aerobic capacity, aerobic power, physical work capacity, and maximal oxygen consumption (VO2max) (Ortega et al., 2014). Similarly, musculoskeletal fitness is synonymous with muscular fitness (Smith et al., 2014).

Historically, physical fitness in children and adolescents was first conceptualized in the late 18th century by physical education teachers and focused primarily on muscular strength and flexibility with the goal of improving 'health fitness'. First assessments of physical fitness were made using performance based tests such as high jump, pull-ups, sprints, standing and running long jump, rope climb, push-ups, and shot put (Dalen et al., 1953). During World War II, these concepts gained traction by incorporating the military's understanding of physical fitness, which traditionally focused on 'motor fitness'. 'Motor fitness' was understood as adequate abilities in a wide range of movement factors, with an emphasis on athletic performance rather than health outcomes (R. R. Pate, 1983). When publications suggested a decline in fitness among American youth[1] (Kraus & Hirschland, 1954), policies were introduced to promote youth health through physical fitness (Hunsicker & Reiff, 1976). As promotion of physical fitness gained importance over time, the understanding of 'being fit' evolved as well. The shift occurred from a broader 'motor fitness' (i.e., cardiorespiratory endurance, muscle strength, body composition, flexibility, agility, strength, speed, balance) to a narrower 'health-related physical fitness' (i.e., cardiorespiratory endurance, muscle strength, body composition, flexibility). The focus herein was on specific components of 'motor fitness' that could be associated with health parameters or disease prevention (R. R. Pate, 1983).

Differentiation between 'motor fitness' and 'health-related physical fitness' over time translated into the distinction between skill-related and health-related components of physical fitness (Blair et al., 1983; Caspersen et al., 1985; Corbin et al., 1978; R. R. Pate, 1983). Following Caspersen et al. (1985), health-related physical fitness includes cardiorespiratory fitness, muscular fitness excluding the muscle strength component, flexibility, and body composition, while skill-related components of physical fitness include the listed subcomponents of motor fitness (see Table 1.1). Of note, the retirement of flexibility as a component of physical fitness has been suggested since it has comparatively little predictive power for meaningful health and performance outcomes in healthy individuals (Nuzzo, 2020). In addition, it is worth noting that a framework for health-related physical fitness proposed by Howley et al. (1986) focused on relative leanness rather than body composition[2] and further incorporated an arousal-relaxation balance, referring to the balance between relaxation (i.e., dominance of the parasympathetic nervous system) and arousal (i.e., dominance of the sympathetic nervous system) in relation to personality traits, external stress, and physical activity[3].

Another differentiation in context of physical fitness occurred between motor skills (‘motorische Fertigkeiten’) and motor abilities (‘motorische Fähigkeiten’). Motor abilities represent latent constructs here, describing general processes of control and functionality that inform and influence the execution of specific movements and motor skills, with motor skills representing specific patterns of movement to accomplish movement tasks (Roth & Willimczik, 1999). Motor abilities (similar to (sub)components of physical fitness) can be divided into (1) conditional abilities such as cardiorespiratory fitness (‘Ausdauer’), muscular strength/fitness (‘Kraft’), and speed (‘Schnelligkeit’), (2) coordinative abilities such as speed (‘Schnelligkeit’) and coordination (‘Koordination’) as well as (3) flexibility (‘Beweglichkeit’) as a passive system for energy transfer. These motor abilities are further separated into [1] aerobic cardiorespiratory fitness (’Aerobe Ausdauer‘), [2] anaerobic cardiorespiratory fitness (’Anaerobe Ausdauer‘), [3] muscular endurance (’Kraftausdauer‘), [4] muscular maximum strength/force (’Maximalkraft‘), [5] muscular speed (’Schnellkraft‘), [6] action speed (’Aktionsschnelligkeit‘), [7] reaction speed (’Reaktionsschnelligkeit‘), [8] coordination (time pressure) (’Koordination (Zeitdruck)’), and [9] coordination (precision) (’Koordination (Präzision)’)[4] (Bös, 2000).

The approach of Bös (2000) assumes that only motor abilities, but not motor skills, can be measured directly. Accordingly, assessment of physical fitness constructs (or, in case of Bös et al. (2000), motor abilities) is done via assessment of motor skills, from which physical fitness information is inferred[5]. Assessment of physical fitness in children and adolescents is usually carried out with field-based tests, as these can easily be administered with little equipment and staff and have good validity and reliability (Artero et al., 2011; Castro-Piñero et al., 2010; Oberger et al., 2006; Ruiz et al., 2011; Safrit, 1990). This allows time-efficient testing of large numbers of subjects (Castro-Piñero et al., 2009; Golle, 2015; Golle et al., 2015).

1.1.1 Maturity and gender

In children, physical fitness is substantially affected by growth and maturation processes that can improve physical fitness irrespective of children’s physical activity (Malina, 2004; Malina & Katzmarzyk, 2006). While growth and maturation are intertwined, growth usually refers to directly measurable changes in body size, weight, proportions and composition, whereas maturation refers to processes that lead the body to a mature (adult) state[6] (Manna, 2014). While all body tissues, organs and organ systems mature at different rates, maturation tends to be independent of chronological age, depending on respective tissues/organs/systems (G. P. Beunen et al., 2006). Skeletal maturity is considered the best single indicator of maturation in children as it is present in all children and is not primarily dependent on maturation of other tissues/organs (Acheson, 1966; G. P. Beunen et al., 2006; G. Beunen & Malina, 1988). Skeletal maturity is usually determined by age at peak height velocity, which refers to the adolescent growth spurt (G. Beunen & Malina, 1988; R. Malina, 1988; Tanner, 1951). The most accurate assessment of skeletal maturity and age at peak height velocity uses radiological examinations of one or more bones (Chumela et al., 1989; Greulich & Pyle, 1959), however, these examinations are expensive, equipment-intensive, require expert interpretation, and carry risks of radiation exposure (Mirwald et al., 2002). Alternatives were proposed by Mirwald et al. (2002), who used longitudinal data from children with predetermined ages at peak height velocity ranging from four years before to three years after peak height velocity. Based on this information, an equation was generated including age, standing height, sitting height and mass, to calculate gender-specific estimates. A similar approach was adopted by Moore et al. (2015), who used a larger data set and created simpler gender-specific equations using only age, standing height, and sitting height to estimate age at maximum height velocity[7].

With regard to gendered determination of maturity, it is important to clarify uses of gender and 'sex' in their relations with each other. It is common practice to classify participants as either male or female under the umbrella term 'sex'[8]. Feminist scholars, scientists, and activists in particular have expressed concern about (mis)uses of these terms over past decades (even centuries), calling for (1) a more nuanced understanding of an individual's biological characteristics (i.e., 'sex') and socially constructed reality in relation to these characteristics (i.e., 'gender'), (2) an awareness of their interrelationships and interdependencies, and (3) a move away from a binary understanding of these two terms (Butler, 1990; Fausto-Sterling, 2020; Johnson & Repta, 2012; R. Jordan-Young & Karkazis, 2019). Classification of 'sex' into two distinct categories based on biological properties does not relate unambiguously to individual biological characteristics. In a scientific context, commonly used properties are chromosomes, gonads, external and internal genitalia, secondary 'sex' characteristics, and hormones, which do not allow for a 'natural' binary definition within and among themselves (Ainsworth, 2015; Blackless et al., 2000; Fausto-Sterling, 2020; R. Jordan-Young & Karkazis, 2019; Oudshoorn, 2005). For example, classification of maturity based on Tanner stages uses inspection of external genitalia in boys and secondary sexual characteristics in boys and girls[9] (Tanner & Whitehouse, 1981). Of note, study on birth rates of intersex people (i.e., people who cannot be identified as "ideally male" or "ideally female" based on their chromosomes, gonads, and internal and external genitalia) found that the birth rate of intersex children is ~1.728% (Blackless et al., 2000), but they are usually considered "not intended by nature" and require surgical "correction" (Fausto-Sterling, 2020). Alternative concepts to a binary 'sex' classification can be found, for example, in the Dominican Republic or in New Guinea, where children with dihydrotestosterone deficiency[10] are recognised as a third 'sex' (G. Herdt, 1990; G. H. Herdt & Davidson, 1988; Rivera-Garza & Herdt, 1996). Additionally, Western scientists such as Fausto-Sterling (Fausto-Sterling & Trajanoski, 2004) have proposed a system of five sexes (i.e., female, female pseudo-hermaphrodite, true hermaphrodite, male pseudo-hermaphrodite and male), still based on external primary genitalia but allowing for more variability and reducing the need for imminent surgical intervention on infants.

In contrast to 'sex', gender is held as continuous construction of specific roles in representation of society (i.e., man and woman as representations for Western societies). This construction occurs in a specific cultural system of meaning in which value is assigned according to social values and hierarchies. Thus, it is linked to political and economic factors in every society/culture (Lauretis, 1987). Gender can be based on individuals identification with an existing gender-identity, in social situations especially in presence of systems of governance however, the construction of gender is still mostly tied to assumptions of a "true" and "binary" 'sex' (Butler, 1990; Fausto-Sterling, 2020; Lugones, 2010; Westbrook & Schilt, 2014). Gender could therefore be based on the individual's identification with a gender identity. However, in social situations, especially in presence of governance systems, construction of gender is currently still mostly tied to a 'true' and 'binary' gender assumption (Butler, 1990; Fausto-Sterling, 2020; Lugones, 2010; Westbrook & Schilt, 2014). In these social situations, the basis of gender assignment[11] may occur in forms such as visual cues in face-to-face interactions, presence of genitalia in sexual situations or even legal cases or political decisions[12], or symbolic reassignment of trans people from one side of the gender binary to the other, which may be required by officials to legally confirm their gender (Westbrook & Schilt, 2014). It is important to note that in processes of colonialism, a binary division of gender into 'men' and 'women' was reserved for colonisers (i.e., bourgeois white Europeans), which also classified them as 'human'. Colonised (i.e., indigenous peoples) were not classified into 'men' and 'women' and thus were classified as ‘non-human’[13] (Lugones, 2010). Accordingly, gendered oppression occurs and intersects with other lines of oppression such as race (Ahmed, 2013; Alcoff, 2006; Collins, 2002; Collins & Bilge, 2016; Fanon, 1986; Gilman, 1994; hooks, 1990; Tate, 2009) or (dis)ability (Clare, 2015; Garland-Thomson, 2011; McRuer & Berube, 2006; Toombs, 2001; Wendell, 1996). Depending on where and how gender is situated[14], lived experiences of gendered individuals as well as access to spaces and resources thus differ depending on social/cultural environments (Lugones, 2010; Westbrook & Schilt, 2014). Some examples of this can be found in income inequality (Mischler, 2021), sports participation (R. Jordan-Young & Karkazis, 2012; Karkazis et al., 2012), or quality of self-reported health (Gómez-Costilla et al., 2022). In line with Kessler et al. (1978), the word ‘gender’ is used instead of ‘sex’ to emphasise social construction of both ‘sex’ and ‘gender’ and to relate gendered research findings to their sociocultural setting rather than defaulting to biological markers of gender differences.

1.1.2 Development of physical fitness

Physiological and morphological changes associated with growth and maturation shape development of children’s physical fitness in different ways and interact with children’s socioecological environment. Several studies have examined this development either through a singular examination of children of different ages (i.e., cross-sectional study) or through several examinations of children over time (i.e., longitudinal study). Ample evidence of improvement in physical fitness during childhood and adolescence (i.e., up to the age of 16-18 years) exists for (1) cardiorespiratory fitness[15] (Albrecht, 2015; Andersen et al., 1976; Castro-Piñero et al., 2011; Catley & Tomkinson, 2013; Fühner et al., 2021; Golle et al., 2014; Miguel-Etayo et al., 2014; Niessner et al., 2020; Oliveira et al., 2014; R. Santos et al., 2014; Tambalis et al., 2016), (2) muscular fitness[16] (Castro-Piñero et al., 2009; Catley & Tomkinson, 2013; Fühner et al., 2021; Golle et al., 2014; Lundgren et al., 2011; Miguel-Etayo et al., 2014; Niessner et al., 2020; Oliveira et al., 2014; Ortega et al., 2023; R. Santos et al., 2014; Tambalis et al., 2016), and (3) speed[17] (Catley & Tomkinson, 2013; Fühner et al., 2021; Golle et al., 2014; Miguel-Etayo et al., 2014; Oliveira et al., 2014; Tambalis et al., 2016). Of note one study remarked that cardiorespiratory fitness plateaued as early as 12 years (Castro-Piñero et al., 2011). (4) Flexibility[18] also shows age-related improvements, however improvements are less pronounced compared to aforementioned physical fitness components (Albrecht, 2015; Catley & Tomkinson, 2013; Golle et al., 2014; Miguel-Etayo et al., 2014; Niessner et al., 2020; Oliveira et al., 2014; R. Santos et al., 2014; Tambalis et al., 2016). (5) Evidence for continuous improvements in balance[19] until adolescence were found in some studies (Lundgren et al., 2011; Miguel-Etayo et al., 2014), while others found that evidence for a stabilisation for dynamic balance around 9 years and for static balance around 12 years (Niessner et al., 2020).

Improvements in physical fitness and especially cardiorespiratory fitness, muscular fitness, and speed also show different gender-related patterns. (1) Cardiorespiratory fitness tests show overall better performances in boys compared to girls (Castro-Piñero et al., 2011; Catley & Tomkinson, 2013; Fühner et al., 2021; Golle et al., 2014; Miguel-Etayo et al., 2014; Oliveira et al., 2014; Tambalis et al., 2016), with more pronounced improvements in boys at age 13 (Niessner et al., 2020; Ortega et al., 2023; R. Santos et al., 2014). Furthermore, studies show that cardiorespiratory fitness peaks in boys at 15-16 years of age (Niessner et al., 2020; Ortega et al., 2023; Tambalis et al., 2016), while in girls it varies between 11-13 (Albrecht, 2015; Ortega et al., 2023) and 14-15 years of age (Niessner et al., 2020; Tambalis et al., 2016). (2) In muscular fitness tests, boys overall tend to outperform girls as well (Castro-Piñero et al., 2009; Catley & Tomkinson, 2013; Golle et al., 2014; Miguel-Etayo et al., 2014; Oliveira et al., 2014; Ortega et al., 2023; R. Santos et al., 2014; Tambalis et al., 2016). However, these differences appear to be comparatively small before puberty, particularly in muscular endurance (Albrecht, 2015; Castro-Piñero et al., 2009; Lundgren et al., 2011; R. Santos et al., 2014), and evidence for gender differences before puberty ranges from studies finding no difference (Castro-Piñero et al., 2009; Lundgren et al., 2011; R. Santos et al., 2014) to boys performing better than girls (Fühner et al., 2021; Golle et al., 2014). Furthermore, longitudinal studies indicate steeper age-related improvements in girls within school grades (Golle et al., 2014), while cross-sectional studies indicate no gendered differences in age-related improvements rates during childhood across grades (Castro-Piñero et al., 2009; Fühner et al., 2021). During and after puberty, muscular strength performance in particular increases in boys and plateaus in girls (Albrecht, 2015; Castro-Piñero et al., 2009; Niessner et al., 2020; Tambalis et al., 2016). (3) Performances in sprint speed tests are better in boys compared to girls (Catley & Tomkinson, 2013; Fühner et al., 2021; Miguel-Etayo et al., 2014; Oliveira et al., 2014; Tambalis et al., 2016), but remain stagnant in all children after the age of 16 (Tambalis et al., 2016). (4) For flexibility and (5) balance performance, most studies found better performance in girls compared to boys (Golle et al., 2014; Miguel-Etayo et al., 2014; Oliveira et al., 2014; R. Santos et al., 2014; Tambalis et al., 2016), while some studies found no evidence of gender differences in terms of balance (Lundgren et al., 2011; Niessner et al., 2020).

In summary, physical fitness improves continuously in childhood up to the age of ~12 years, with subsequent different trajectories in boys and girls, such as a greater increase in cardiorespiratory fitness in boys (Niessner et al., 2020; R. Santos et al., 2014) compared to a possible plateau in girls (Albrecht, 2015; Castro-Piñero et al., 2011). For a more detailed presentation of development curves of physical fitness components in children, see Albrecht et al. (2015), Niessner et al. (2020), or Ortega et al. (2023). Physical, physiological, and biological changes related to growth and maturation are often cited as the reason for difference in development of physical fitness with age and between genders (Albrecht, 2015; Castro-Piñero et al., 2011; R. Santos et al., 2014; Tambalis et al., 2016). Due to the complexity of maturation and growth (G. P. Beunen et al., 2006), effects of specific maturation processes on physical fitness performances in their interrelationships are still unclear.

1.1.3 Secular trends of physical fitness

Aside from interactions with growth and maturity, physical fitness in children of same age and location changes over time, as so-called secular changes. Particularly, concerns of declining physical fitness among youth has gained much attention in recent decades and has led to research assessing these trends at an local, national, and international level (Dooley et al., 2020; Fühner et al., 2020; Tomkinson et al., 2003, 2019, 2020; Tomkinson & Olds, 2007).

1.1.3.1 International secular trends

International[20] analyses of secular trends in children aged 6 to 19 years were first published from the early 2000s and have since increased in popularity. Cardiorespiratory fitness[21] showed an overall improvement from 1958 to 1970 (Tomkinson & Olds, 2007), followed by a decline until 2000 ~ 2010 (Fühner et al., 2020; Tomkinson et al., 2003, 2019; Tomkinson & Olds, 2007), where it stabilised and reached a floor (Fühner et al., 2020; Tomkinson et al., 2019). Regarding muscular fitness, international trends are available for the subcomponents muscular strength, muscular power, and muscular endurance. International secular trends for muscular strength[22] yielded conflicting results, showing an overall small negative quadratic trend (Fühner et al., 2020) in contrast to overall progressive improvements until 2017 (Dooley et al., 2020). For muscular power[23], international secular trends showed overall improvements followed by subsequent declines with peaks in ~1985 (Tomkinson, 2007) or ~2000 (Tomkinson et al., 2020). For muscular endurance[24] large international improvements with a slowdown that stabilised around 2010 and declined thereafter were found (Kaster et al., 2020). Studies assessing international secular trends for speed[25] found improvements up to ~1970, followed by steady declines (Tomkinson, 2007), while another study found slight to moderate improvements from 1970 to ~2015 (Fühner et al., 2020). Contradictory results for muscular strength and speed could be at least partly explained by differences in inclusion criteria between studies. For example, Fühner et al. (2020) used a comprehensive approach to muscular strength that allowed different tests to be included in the analysis, while Dooley et al. (2020) focused on handgrip strength only.

Explanations of varying trend directions and developments can be combined into an interrelated model of physiological/psychological, physical, behavioural, and social factors (Tomkinson, 2004). Following this model, several subanalyses have been conducted to further identify and elucidate moderator variables of these trends. Estimating trends for larger geographical regions[26], as well as individual countries showed trends in geographic regions to be similar to international trends (Tomkinson, 2007; Tomkinson & Olds, 2007), while greater variability is shown around directions of national trends regarding direction and uniformity (i.e., linear or curvilinear) (Dooley et al., 2020; Kaster et al., 2020; Tomkinson et al., 2003, 2019, 2020). For gender, most studies found similar trends for boys and girls (Dooley et al., 2020; Kaster et al., 2020; Tomkinson et al., 2003, 2019, 2020; Tomkinson, 2007), with the exception of cardiorespiratory fitness, where a steeper decline was found for boys (Fühner et al., 2020; Tomkinson & Olds, 2007). Correlating international secular physical fitness trends with international secular trends of body mass index[27] (BMI) found a positive correlation for handgrip strength in children but not adolescents, with higher body mass indices being associated with stronger handgrips (Dooley et al., 2020), but found no associations with shuttle run, standing long jump, and sit up performances (Dooley et al., 2020; Kaster et al., 2020; Tomkinson et al., 2019, 2020). Correlating international physical fitness trends with international trends in moderate to vigorous and vigorous physical activity[28] revealed a positive correlation between sit up performances and vigorous physical activity, indicating that being vigorously physically active for one hour four times a week might be beneficial for secular trends in sit up performance or vice versa (Kaster et al., 2020). No significant correlations were found between handgrip strength, standing long jump, and shuttle run performance with physical activity (Dooley et al., 2020; Tomkinson et al., 2019, 2020). Comparing secular trends to developments in socioeconomic variables at country level such as gross national income per capita[29] (Tomkinson, 2007; Tomkinson & Olds, 2007), Gini index[30] (Dooley et al., 2020; Kaster et al., 2020; Tomkinson et al., 2019, 2020), human development index[31] (Dooley et al., 2020; Kaster et al., 2020; Tomkinson et al., 2019, 2020), and urbanisation rate[32] (Dooley et al., 2020; Kaster et al., 2020; Tomkinson et al., 2019, 2020) found (1) trends in speed and muscle strength to be favourable in low income countries but detrimental in high income countries between 1970 and 2000 (Tomkinson, 2007), (2) smaller declines in shuttle run performance to be associated with a larger increase in gross domestic product (Tomkinson et al., 2003), (3) correlations between Gini index with shuttle run performances, with countries approaching income equality showing more favourable trends (Tomkinson et al., 2019), and (4) no significant associations for the remaining comparisons/correlations.

In summary, secular trends of physical fitness show a complex development with ambiguous directions for different components of physical fitness, as well as interdependencies with the social world. This emphasises the need for an understanding of these interdependencies in order to properly contextualise secular fitness trends and be able to create meaningful and effective interventions.

1.1.3.2 Regional secular trends and interrelations - EMOTIKON study

Due to the regional relevance for this thesis, regional secular trends for the federal state of Brandenburg, Germany, are presented. These trends were assessed in the EMOTIKON study, which also served as foundation for the SMaRTER study (see Section 2.3). Measurements of physical fitness are taken using a fitness battery of six tests consisting of 6 min run, standing long jump, ball push test, 20 m sprint, one leg balance test, and star run (Teich, 2023). A detailed descriptions of these tests can be found in Section 2.4.2.1.

Secular trends in the EMOTIKON study were analysed for five of the six tests mentioned above (one leg balance test was not included in the analysis) from 2011 to 2019 in 108,295 children aged 8 to 8.99 years spread across 515 schools. Of note, in this analysis, contrasts were specified to compare performance trends in (1) star run to 6 min run, (2) 20 m sprint to star run, (3) standing long jump to 20 m sprint, and (4) standing long jump to ball push test. Analyses showed linear increases in test performances with chronological age and better performances of boys compared to girls in all included fitness tests. Tests differed in magnitude of respective gender and age effects. Interestingly, no evidence of an age-gender interaction was found for any of the tests or contrasts, despite an abundance of statistical power. Evidence for secular trends was only found for 6 min run with declines and for 20 m sprint with improvements in test performances. Additional findings were that 6 min run, 20 m sprint, star run and standing long jump tests form a correlated construct representing a potential conception of 'physical fitness', and that 'physically fit schools' show greater age-related improvement (Fühner et al., 2021).

In another publication, Fühner et al. (2022) analysed physical fitness of children depending on their age at school enrollment (i.e., 30 September). Three different groups were defined as (1) keyage (108,296 children aged 8 to 8.99 years; sample from Fühner et al. (2021)), (2) younger than keyage (2,586 children aged 7 to 7.99 years), and (3) older than keyage (26,540 children aged 9 to 9.99 years). Their analyses showed that keyage children performed better than older than keyage children, especially in the star run test, and that gender differences were present but reversed for older than keyage children. In addition, analyses showed that younger than keyage children performed better than keyage children, especially in the star run test. The authors hypothesise that differences between groups are related to differences in maturation that lead to children being enrolled in school earlier or later (Fühner et al., 2022).

In unpublished analyses on sociocultural, -economic, and -structural interrelations of physical fitness in the federal state of Brandenburg, the EMOTIKON research team investigated effects related to proximity to Berlin[33]. They found better performance in 6 min run, 20 m sprint, and standing long jump in children who lived closer to Berlin, although there was no evidence for the effect of proximity to Berlin on the direction of secular trends in the analysed tests (R. Kliegl & Teich, 2022). A further analysis linking the EMOTIKON data set to results of the 2021 federal election in the state of Brandenburg, Germany found a negative correlation between 6 min run performance and percentage of right-wing members in local councils and a positive correlation between performance in 6 min run performance and percentage of Green and Liberal party members in local councils. These effects were reversed for ball push test performances (R. Kliegl & Teich, 2022). This means, for example, that areas with a high proportion of right-wing party members and a low proportion of green and liberal party members in local councils have comparably lower 6 min run performances and higher ball push test performances. With regard to the tendency of low-income voters to prefer right-wing parties and of higher-income voters to prefer green and liberal parties (DIW, 2017), these results emphasise the sociocultural, -economic, and -structural conditions of physical fitness and its health-related correlates.

As regional secular trends in 6 min run and 20 m sprint (Fühner et al., 2021) share parallels with international trends in cardiorespiratory fitness (Fühner et al., 2021; Tomkinson et al., 2019; Tomkinson & Olds, 2007) and speed (Fühner et al., 2020), these trends might be caused by a similar component. Particularly striking is the relation to different aspects of income in its specific sociocultural context, such as the Gini-index (Tomkinson et al., 2019) or the association between income and electoral behaviour (DIW, 2017; R. Kliegl & Teich, 2022).

1.2 Physical fitness and weight status

Over the past few decades, considerable evidence has been found to show low physical fitness to be a risk factor associated with a wide range of health problems in children. For example, low muscular fitness in childhood and adolescence is associated with an increased risk of cardiovascular disease, an increase in metabolic risk factors and an increase in obesity, while high muscular fitness is associated with improved bone health (García-Hermoso et al., 2019; Ramírez-Vélez et al., 2016; Smith et al., 2014). Similarly, low cardiorespiratory fitness is associated with an increase in cardiovascular risk factors such as abnormal blood lipids, hypertension, and increased obesity (Ruiz et al., 2009), as well as symptoms of depression (Esmaeilzadeh, 2014, 2015). While there are a variety of different negative health effects, the following sections primarily focuses on associations between physical fitness and weight status.

Weight status in mostly used in physical fitness research to determine children with overweight/obesity, which can vary depending on its definition. The most commonly used tool for assessing overweight/obesity is the BMI (i.e., mass/height^2). For adults, BMI cut-off scores for overweight are 25 kg/m^2 and for obesity 30 kg/m^2, which are predominantly used for adults in Western countries (Akram et al., 2000). For children and adolescents, age- and gender-specific BMI cut-off values have been estimated based on adult prevalence to account for maturation and growth-related changes in the body (Cole et al., 2000; Cole & Lobstein, 2012; Kromeyer-Hauschild et al., 2001). BMI is most commonly used in population-based studies as it is easily accessible and provides a reasonable estimate of prevalence of overweight and obesity in a larger population (Akram et al., 2000; Bentham et al., 2017; Ng et al., 2014; Y. Wang & Lobstein, 2006). It is assumed that changes in body mass (which affect the ratio of height to body mass) are mainly due to an increase in body fat (Akram et al., 2000; Cole et al., 2000; Nishida et al., 2004). However, for exercise-based interventions, the general goal is to reduce fat mass and increase muscle mass. The BMI is not suitable for differentiating between fat and muscle mass, thus other measures and methods, such as dual-energy X-ray absorptiometry, skinfold thickness, waist circumference, or bioelectrical impedance analysis, can be used (Lindsay et al., 2001; Lobstein et al., 2004; Steinberger et al., 2005; Tyrrell et al., 2001; Wohlfahrt-Veje et al., 2014).

In recent decades, both scientists and health-oriented institutions have expressed concern about growing numbers of overweight and obese people worldwide. A survey assessing international changes of overweight and obesity prevalences[34] in children and adolescence aged 2 to 19 years in 188 countries found an overall increases between 1980 and 2013. In developed countries[35] prevalences increased from 16.2% to 22.6% in girls and from 16.9% to 23.8% in boys, and for developing countries from 8.4% to 13.4% in girls and from 8.1% to 12.9% in boys (Ng et al., 2014). Another study examined secular trends in mean BMI among 24.1 million children and adolescents aged 5 to 17 years from 1975 to 2016. Across all age groups, an international increase from ~17.2 kg/m^2 to ~18.6 kg/m^2 was found for girls (i.e., ~0.32 kg/m^2 per decade), and from ~16.8 kg/m^2 to ~18.5 kg/m^2 for boys (i.e., ~0.4 kg/m^2 per decade). Regarding prevalence in age standardised obesity[36], found increases from 0.7% to 5.6% in girls and from

0.9% to 7.8% in boys from 1975 to 2016 (Bentham et al., 2017). A study that did not utilise BMI examined secular changes in skinfold thickness and percent body fat in 458 547 children and adolescents aged 0 to 18 years from 1951 to 2003 in 30 industrialised countries[37]. The analyses showed an increase in skinfold thickness of ~0.49 mm and percentage body fat of ~0.86% per decade with highest rates in children age 10 to 14 years (Olds, 2009).

These changes are of particular concern to public health officials, as overweight and obesity are associated with negative health outcomes. For example, childhood overweight or obesity[38] is associated with increased risk factors for cardiovascular disease such as elevated trigylceride, LDL, and HDL cholesterol levels, insulin, and elevated systolic and diastolic blood pressure (Freedman et al., 2009, 2013). Moreover, a meta-analysis examining diagnosis and symptoms of depression in overweight and obese[39] children (0 to 12 years) and adolescents (13 to 21 years) showed that obese children have a higher risk of developing depression (OR = 1.34) (Quek et al., 2017). While children who are classified as overweight or obese face immediate risks to their physical and mental health, several studies have examined links between childhood overweight and obesity classifications and health risks later in life. For example, one study found an increased risk of death from coronary heart disease, atherosclerosis, colorectal cancer, and gout in 256 men who were classified as overweight in their youth[40] (Must et al., 1992). A systematic review examining effects of obesity in childhood (i.e., 2 to 12 years) and adolescence (i.e., 13 to 19 years) on morbidity and mortality in adulthood (i.e., ≥ 19 years) found strong evidence of increased risk of type 2 diabetes, hypertension, coronary heart disease, colorectal cancer, and all-cause mortality[41] (Park et al., 2012). In addition, several studies have shown that childhood and adolescent overweight/obesity continues into adulthood (Singh et al., 2008), and that the number of years living with obesity further increases all-cause mortality, cardiovascular mortality, cancer mortality, and mortality from other causes (Abdullah et al., 2011).

Due to negative health effects associated with overweight and obesity in childhood and adolescence, the relationship between physical fitness and weight status is considered a possible cause and explanation for negative health effects associated with low physical fitness. However, the relationship between physical fitness and weight status is ambivalent, multidirectional, and varies for different components of fitness and assessments of obesity. For example, in the aforementioned studies assessing international secular trends in physical fitness, no overall significant national-level correlations with secular trends in BMI were found for different components (Dooley et al., 2020; Kaster et al., 2020; Tomkinson et al., 2019, 2020) (see Section 1.1.3.1 for more details). Only subanalyses of secular trends in BMI and handgrip strength for children and adolescents seperately showed a positive correlation for children (r = .55, 95%CI = .03 to .84) but not for adolescents (r = .21, 95%CI = -.36 to .67). This may indicate that positive effects of BMI on handgrip strength may lessen with age. As a reason for the positive direction of the relationship, the authors hypothesised that improvements in handgrip strength are due to an increase in lean muscle mass, which is part of increases in BMI (Dooley et al., 2020). Similarly, in a study examining the relationship between handgrip strength and BMI or percentage body fat in children[42], found a positive correlation between BMI and handgrip strength, while the correlation between body fat and handgrip strength was negative[43] (Sartorio et al., 2002). In other tests of muscular fitness, the relationship of weight status with physical fitness changes direction. For example, in a study examining the relationship between muscular strength and weight status in American children and adolescents, a positive relationship was found between weight status with knee extension and handgrip strength, with higher performance associated with higher weight status. However, while in plank and pull-up tests, there was an inverse relationship, with an increase in weight status associated with a decrease in performance[44] (Ervin et al., 2014). It has been theorised that muscle strength in tests that require lifting body mass is inversely related to weight status, as performances are more affected by body mass and especially body fat mass, while tests that focus on specific muscles without utilising total body mass show a parallel relationship due to better utilisation of lean body mass (Deforche et al., 2003; Ervin et al., 2014). Similarly, cardiorespiratory fitness, particularly in tests measuring running performance over a given distance or time, is consistently better in normal-weight children compared to overweight and obese children (Abdelkarim et al., 2020; Palomäki et al., 2015). For example, a study examining the relationship between weight status, 20 m shuttle run performance, and physical activity in adolescents aged 15-16 years showed that cardiorespiratory fitness was better in normal-weight adolescents than in overweight adolescents, regardless of physical activity and gender[45] (Palomäki et al., 2015). Another study assessing effects of weight status on 9 min run performance in 519 Brazilian students could show better performances in normalweight compared to overweight, and overweight compared to obese children[46]. They were also able to show that weight status had a greater impact on 9 min run performance compared to other fitness tests such as sit-and-reach, stationary long jump, 1 min curl-up, modified pull-up, medicine ball throw, 20 m run, and 4 m shuttle run (Dumith et al., 2010).

In relation to aforementioned interrelations between physical fitness and socioeconomic variables, it is important to highlight socioeconomic correlates shaping weight status in youth. For example, a systematic review of 158 papers examining associations between socioeconomic position and weight status among adolescents in the 21 richest countries between 1990 and 2005 found an inverse association in 60.7% of the studies[47]. 18.7% found no association and 20.9% found the association to be confounded by other variables such as age, gender, or ethnic group, and only 1.1% found a parallel association (Barriuso et al., 2015). Similar results were obtained in a study examining effects of physical activity, television viewing, video game play, socioeconomic status, and ethnicity on BMI of American adolescents aged 10 to 16 years[48]. The most striking associations were found for ethnicity, socioeconomic status, and gender, with higher rates of overweight in African American children, low socioeconomic status, or girls compared to White children, high socioeconomic status, or boys (McMurray et al., 2000). These findings suggest that prevalence of overweight and obesity runs along lines of societal marginalisation, which would support an intersectional approach[49]. A study assessing obesity prevalence and its social correlates in Black and White American adolescents aged 15.2 ± 1.6 years attending public school found that Black girls comparatively assigned themselves to the lowest social stratification and had the highest BMI. Analysis of socioeconomic status revealed that lower-income and Black students were more likely to be overweight and that, on average, Black students came from lower-income families compared to White students[50] (Goodman et al., 2003). Interestingly, a study that examined trends in obesity prevalence among 534 children aged 2 to 4 years over a three-year period in relation to household income[51] found that obesity prevalence increased at a greater annual rate in low-income households compared to higher-income households, highlighting longevity of income-related differences in obesity prevalence (Kunin-Batson et al., 2023).

In summary, the relation on weight status varies between and within different physical fitness components in its magnitude as well as direction depending on different test properties. Further, as weight status is associated with specific socioeconomic markers, interpretations of effects of weight status on physical fitness need to consider its socioeconomic context.

1.3 Physical fitness and executive function

In recent decades, in context of physical fitness, cognitive performance, particularly executive function[52], has received increasing attention in children and adolescents (Chu et al., 2019;

Donnelly et al., 2016; Ruiz-Ariza et al., 2017). Executive functions describe top-down control processes against automatic or instinctive decisions or actions, and is subdivided into the domains inhibition, working memory and cognitive flexibility (Diamond, 2013; Haapala, 2013). Inhibitions is at the core of executive function and is ascribed to the control of attention, actions and reactions, while working memory describes the ability to hold and actively process information (Diamond, 2013; Haapala, 2013). Cognitive flexibility is found at the intersection of inhibition and working memory and is described as changing perspective on a problem or adjusting to new rules (Diamond, 2013) Studies examining the relationship of physical fitness with executive function have found a small but positive association. For example, in studies comparing high-fit with low-fit children on basis of their cardiorespiratory fitness, high-fit children were found to have better allocation of attentional and working memory resources (Hillman et al., 2005) and better response accuracy (Hillman et al., 2009) compared to low-fit children[53]. A study examining the relationship between shuttle run performance and executive functions in children aged 12 to 15 years found a small but significant positive correlation between shuttle run performance and cognitive flexibility, but not with problem-solving skills[54] (Niet et al., 2014). Another study that examined the association between a 10-minute interval running performance and a composite of executive functions in 378 children (aged 9 to 10 years) found a small but significant positive association[55] (Kvalø et al., 2019). Skog et al. (2020) examined the relationship between maximal oxygen uptake and maximal power output with working memory, short-term memory, visual learning and memory, paired associated learning, attention, reaction time, and 'executive function' and found that maximal oxygen uptake was positively associated with working memory, visual learning, and associated learning[56]. In contrast, Havercamp et al. (2021) found no relationship between performance in 20 m shuttle run and visuo-spatial working memory, verbal working memory, and interference control in adolescents[57]. Positive direction of the relationship between cardiorespiratory fitness and executive function is attributed to cardiorespiratory fitness-induced structural changes in the brain, such as synaptic plasticity (Chaddock et al., 2011), grey matter development (Cotman et al., 2007), white matter integrity (Chaddock-Heyman et al., 2014), activated growth factors (Cotman et al., 2007; Hillman et al., 2008), and/or improved cerebral blood flow (Chaddock et al., 2012; Tyndall et al., 2018).

Studies examining the relationship between muscular fitness and executive functions in children do not show such a clear pattern. In studies investigating the relationship between physical fitness and executive functions, tests assessing muscular fitness parameters usually calculate a composite score with a test assessing cardiorespiratory fitness to achieve a proxy score for physical fitness (Chu et al., 2019). In these studies, similar correlations are found as in aforementioned studies on cardiorespiratory fitness. For example, using a composite score consisting of 20 m shuttle run and maximum ball throwing performance, Marchetti et al. (2015) found a slight positive relationship between the composite score for physical fitness and working memory updating and inhibition in adolescents[58]. Only a few studies have directly measured muscle strength, and they have not shown a consistent pattern linking muscular fitness and executive function. Van der Niet et al. (2014) found significant small positive associations between performance in sit-up, 10 x 5 m shuttle run, and standing long jump with cognitive flexibility but not with problem-solving skills[59]. Whereas Haverkamp et al. (2021) found no correlation between standing long jump/sit up performance and any included executive function test[60]. Weak but positive associations between muscular fitness and executive function could be due to strength training-induced increases in insulin-like growth factor 1 found in older individuals (Cassilhas et al., 2007). Assuming that this relationship is transferable to children and adolescents, increases in neuronal growth could at least partially explain positive relations between muscular fitness and executive function.

In addition to cardiorespiratory and muscular fitness, fitness tests that include a cognitive component were assessed in relation to executive function. Haverkamp et al. (2021) found a positive association between 10 × 5 m shuttle run and plate tapping performances with visuospatial working memory, information processing and control, and interference control[61], Van der Niet et al. (2014) found significant small positive associations between 10 x 5 m shuttle run performance and cognitive flexibility[62], and Marchetti et al. (2015) found a positive association between inhibition and performance in pendulum running in high fitness children but not in low fitness children[63]. These positive relationships are attributed to neural network activations as a common denominator associated with both motor and executive functions (Diamond, 2006; Koziol et al., 2014).

1.4 Physical education interventions

In context of some negative secular trends internationally (see Section 1.1.3.1) and nationally (see Section 1.1.3.2) and the correlation between physical fitness and adverse health outcomes (see Section 1.2), researchers have called for interventions to counter secular trends and promote children's physical fitness and health (Fühner et al., 2020, 2021; Tomkinson et al., 2020). A common approach is to implement physical fitness interventions in schools, where almost all children can be reached[64] (Kriemler et al., 2011). In addition, schools provide a convenient environment for implementation of physical fitness interventions, as they have access to trained instructors and training facilities and equipment. A meta-analysis evaluating the impact of qualitative (i.e., incorporating fitness exercises or teaching strategies into existing physical education) and quantitative (i.e., increasing the number of physical education lessons per week) physical education interventions found that both interventions have a positive impact on physical fitness, with quantitative physical education interventions having a greater impact on components of physical fitness compared to qualitative interventions (García-Hermoso et al., 2020). Since quantitative physical education interventions seem to be effective in improving physical fitness, the following sections highlight effects of these interventions on components of physical fitness as well as on parameters of body composition and cognition.

1.4.1 Physical education interventions and physical fitness

Several studies have investigated effects of additional physical education on physical fitness. Ardoy et al. (2011) analysed effects of two hours of standard physical education, four hours of standard physical education, and four hours of intensive physical education in 67 Spanish adolescents (aged 12-14 years; 35.8% girls) over a 16-week period on different components of physical fitness[65]. They found larger increases in 20 m shuttle run and sit and reach performances in both intervention groups compared to a control group, and in 4 x 10 m shuttle run performance in the intensive physical education group compared to the control group. No differences were found between groups regarding standing long jump performances. A study evaluating effects of a comprehensive physical education intervention that involved doubling weekly physical education time (i.e., from 90 min/week to 180 min/week) over a three-year period in 696 preschool children (aged 6-7 years at baseline)[66] found no differences in VO2 peak between intervention and control groups (Bugge et al., 2012). A similar study involving 503 first and fifth grade children compared three hours of regular physical education with five hours of special physical education over the course of a school year. They found small differences in shuttle run performances between intervention and control groups, favouring the intervention group[67] (Kriemler et al., 2010). Löfgren et al. (2013) investigated daily physical education lessons (i.e., 5 x 40 min./week) compared to regular physical education lessons (i.e., 2 x 30 min./week) in 232 children aged 7 to 9 years at baseline over a 2-year period[68]. They found better knee extension strength at 180° in boys and girls in the intervention group compared to controls, better intervention-related knee flexion strength at 180° in boys only, and better vertical jump performance in girls in the intervention group compared to girls in the control group. However, in a study with a similar design involving 189 elementary school children, girls in the intervention group with daily physical education (i.e., 5 x 45 minutes/week) were found to have higher push-up and curl-up performances compared to girls in the control group with regular physical education (i.e., 1 x 45 minutes/week). No differences were found for boys overall or for either gender in shuttle run performance[69] (Reed et al., 2013).

1.4.2 Physical education interventions and weight status

Several of the aforementioned interventions also included health-related anthropometric parameters such as total skinfolds (Ardoy et al., 2011; Bugge et al., 2012; Kriemler et al., 2010), percent body fat (Ardoy et al., 2011; Löfgren et al., 2013), other body fat distribution proxies[70] (Ardoy et al., 2011; Bugge et al., 2012; Kriemler et al., 2010; Löfgren et al., 2013; Reed et al., 2013), and specific cardiovascular disease risk factors[71] (Bugge et al., 2012; Kriemler et al., 2010) as secondary outcomes. While Kriemler et al. (2010) and Reed et al. (2013) found evidence of a reduction in body fat distribution in favor of children participating in the interventions (i.e., skinfold thickness, BMI), Löfgren et al. (2013) found evidence of an intervention-related increase in body fat, with more favorable trends in the control condition[72]. Ardoy et al. (2011) and Bugge et al. (2012) found no evidence of intervention effects with respect to body fat distribution. However, with respect to cardiovascular disease risk factors, there was evidence of more favorable trends associated with intervention participation (Bugge et al., 2012; Kriemler et al., 2010). Another study examined effects of four additional hours of physical education on overweight and obesity parameters as primary outcomes in 632 children aged 8 to 13 years[73]. No differences were found in percent body fat and BMI between six lessons of physical education and two lessons of physical education. However, a significant effect was found on prevalence of overweight and obesity, with children who participated in the intervention being at lower risk for overweight and obesity (Klakk et al., 2013).

1.4.3 Physical education interventions and executive function

Few studies have examined effects of additional physical education on executive function or similar cognitive domains in children and adolescents. Of the aforementioned studies, only Reed et al. (2013) included cognitive measures such as puzzle-solving ability and perceptual speed[74]. While no clear intervention-related effects were found for puzzle-solving ability, greater improvements in perceptual speed were found for girls in the intervention group compared to girls in the control group. Another study examining effects of two weekly additional cognitively enhanced tennis lessons compared to one weekly lesson of normal physical education on inhibition and working memory in 9 to 10 year old children found intervention-related improvements in inhibition, but not memory, and only in overweight children[75] (Crova et al., 2014). However, several studies examined effects of quality/intensity of physical education on executive function in children and adolescents and found ambiguous evidence of the relationship between intensity of physical education and cognition. For example, a study examining effects of high-intensity interval training on memory, selective attention, and concentration in adolescents aged 13.7 ± 1.3 years found that selective attention and concentration, but not memory, increased more in the intervention group compared with the control groups that received regular physical education[76]. Conducting additional subanalyses, the authors found that these effects were more pronounced in more inactive children (Martínez-Lopez et al., 2018). In contrast, a study examining effects of physical education combined with either aerobic high-intensity interval training or resistance and aerobic high-intensity interval training in adolescents aged 14 to 16 years found no effect of intensive physical education on trail making test performance[77] (Costigan et al., 2016). A meta-analysis examining effects of physical activity on executive functions in children and adolescents found an overall improvement with physical activity, regardless of age. Regarding study design, they found that curricular exercises and extracurricular programs were effective, while physical activity integrated into school curriculums did not significantly improve executive functions. They were also able to show that fewer than 5 sessions per week seem to be most effective with a session duration of less than 90 minutes (Xue et al., 2019).

1.4.4 Overall effectiveness of physical education interventions

In summary, additional physical education can have a positive impact on physical fitness, executive functions, and weight status of children and adolescents. However, these advantages seem to be mitigated by gender and physical condition of participating children and young people, as well as by duration, intensity, and objective of additional physical education. While some studies have examined effects of additional physical education in overweight children (Crova et al., 2014; Klakk et al., 2013), there are no studies to date that examine effects of additional physical education in children with low levels/deficits in physical fitness. Improving physical fitness in children with physical fitness deficits is important because these children are at higher risk of adverse health effects and might also progress more in physical fitness due to their lower baseline levels. Accordingly, further research is needed to investigate effects of additional physical education on physical fitness, weight status, and executive functions in children with physical fitness deficits.

1.5 Physical fitness, health, and cognition during the Covid 19 pandemic

As this thesis was written in late 2022, it is important to consider the impact of the global Covid 19 pandemic and its implications for children and adolescents. To prevent spread of Sars-CoV-2, governments have enacted regulations restricting access to areas of direct human interaction such as workplaces, schools and public spaces. Due to geographical location as well as thematic setting of the intervention, on which this thesis is focused, measures to contain spread of Sars-CoV-2 in schools and organised sports are described for the federal state of Brandenburg, Germany. From March 18th, 2020 to April 20th, 2020, schools were completely closed (MBJS, 2020c, 2020a) and slowly reopened, using models where children could attend school two to three times per week, following hygiene protocols (e.g., regular testing, social distancing, wearing masks) (MBJS, 2020b). From August 9th, 2021 onwards, schools were fully open (MBJS, 2021a) and kept open until winter 2021/22, with compulsory school attendance removed to allow schools to respond to local infection prevalences (MBJS, 2022a, 2022b). Compulsory school attendance was reintroduced from March 7th, 2022 (MBJS, 2022b). During reopening of schools from 2020 to 2021, physical education classes were conducted only outdoors following social distancing and other hygiene measures (MBJS, 2020a) and from August 9th, 2021, physical education classes were conducted without any restrictions (MBJS, 2021b). A more detailed chronological list of regulations for school operations can be found elsewhere[78]. Organised sports were suspended from March 18th, 2020 to May 15th, 2020 (MBJS, 2020c, 2020a) and were subsequently only allowed outdoors, in small groups following social distancing protocols (MBJS, 2020a). From May 28th, 2020, organised indoor activities were allowed following social distancing procedures and a facility-specific hygiene concept (MBJS, 2020b). Operation under 2G regulations (i.e., being vaccinated/cured or tested) was reintroduced on November 24th, 2021 (MBJS, 2021c).

Considering that severity of lockdown measures directly affects children's physical activity (Kharel et al., 2022; Paterson et al., 2021; S. C. E. Schmidt et al., 2020, 2022), the question arose how these lockdown measures affected children's physical fitness. Teich et al. (2023) analysed effects of policies described above on secular development of physical fitness in 107,558 third grade children in the federal state of Brandenburg, Germany, using data from the EMOTIKON study (see Section 1.1.3.2). In 87,395 key-age children[79] they found a significant negative impact of covid pandemic on a composite physical fitness score and on 6 min run, 20 m sprint, and star run performance. There was a significant positive trend in standing long jump performance but no significant effect in ball push test and single leg balance performance. In particular, negative impact in running tests resulted in annual development costs/delays of 3.5 months for 6 min run and star run performance and 2.1 months for 20 m sprint performance. The negative trend in composite physical fitness score resulted in an annual development cost/delay of 1.4 months, while annual developmental gain in standing long jump performance was 1.6 months. Secondary analyses for 22,761 children older than keyage showed negative covid pandemic effects on 6 min run, star run, ball push test, and one-leg balance test performance, whereas 1,321 children younger than keyage showed no significant covid pandemic effects. Similar trends were found in a study assessing the impact of the Covid pandemic in Slovenian 6th and 8th grade children[80], who showed declines in all tests except hand tapping test and found greatest declines in physical fitness index, 600 m run, polygon course backwards, bent-arm hang, and 60 m sprint performance. They were also able to show that declines were greater for children in rural areas compared to urban children (Pajek, 2022).

While there are negative trends in physical fitness associated with lockdown measures, these measures were put in place to mitigate health risks associated with SARS-CoV-2 infections. Because of the novelty of the virus, knowledge of effects of SARS-CoV-2 infection in general and in children in particular is still sparse and requires further investigation, especially with regard to long-term effects. Acute effects of SARS-CoV-2 infection appear to be less severe in children and adolescents aged 0-18 years than in adults (Liguoro et al., 2020; Mehta et al., 2020), with ~42.5% of children with SARS-CoV-2 infection experiencing mild symptoms and 39.6% experiencing moderate symptoms. ~2% of children with SARS-CoV-2 infections required treatment in a pediatric intensive care unit, and estimated overall mortality rate was <.1%[81] (Liguoro et al., 2020). However, apart from acute SARS-CoV-2 infections, several long-term effects have been observed in children, such as multisystem inflammatory syndrome (Brodin, 2022; Molloy et al., 2022; Nygaard et al., 2022) or post-COVID-19 disease. Multisystem inflammatory syndrome in children causes fever, hypotension/shock, myocardial dysfunction/coronary abnormalities, and elevated inflammatory markers, amongst other symptoms, and can occur 1 to 2 months after asymptomatic or paucisymptomatic SARS-CoV-2 infection (Brodin, 2022; Molloy et al., 2022). Although prevalences of multisystemic inflammatory syndrome in children after SARS-CoV2 infection is comparatively low (i.e., 51 of 175 458 unvaccinated individuals aged 0 to 18 years), prevalences vary by SARS-CoV-2 variant and underlying mechanisms are not fully understood (Nygaard et al., 2022). The term post-COVID-19 condition refers to symptoms lasting several months after an acute infection, such as fatigue, shortness of breath, insomnia, difficulty breathing, nasal congestion, difficulty concentrating, muscle pain, exercise intolerance, weakness, and walking intolerance (Izquierdo-Pujol et al., 2022; Soriano et al., 2022; Zimmermann et al., 2021). Overall prevalences in children and adolescents are unclear, and early estimates ranged from 4% to 66% (Zimmermann et al., 2021) to 1% to 30% (Izquierdo-Pujol et al., 2022), depending on included symptoms. These findings gain importance as a recent study found that severity and persistence of acute and long-term symptoms increase with the number of SARS-CoV-2 infections (Bowe et al., 2022).

Accordingly, SARS-CoV-2 infections in children pose a significant health risk with as yet unknown long-term consequences. Whether these risks outweigh negative effects of lockdowns and social distancing measures on for example physical activity and physical fitness (Kharel et al., 2022; Pajek, 2022; Paterson et al., 2021; Teich et al., 2023) cannot be conclusively assessed. Full opening of schools (MBJS, 2021a), reinstatement of compulsory school attendance (MBJS, 2022b), and lack of comprehensive testing in schools from April 30th, 2022 (MBJS, 2022c) combined with a high prevalence of asymptomatic acute infections (Liguoro et al., 2020; Mehta et al., 2020) but potentially serious longterm consequences in children (Bowe et al., 2022; Izquierdo-Pujol et al., 2022; Nygaard et al., 2022; Zimmermann et al., 2021), however, seem quite concerning.

1.6 Main hypothesis

Based on the outlined need for further research to assess effects of physical education on physical fitness, weight status, and executive function in children with deficits in their physical fitness (see Section 1.4.4), a study (i.e., SMaRTER study) was conducted guided by the following primary working hypothesis:

> *Children with physical fitness deficits show greater improvements in physical fitness and cognitive performance following participation in a remedial physical education program compared to a control condition.*

The SMaRTER study and its sample are described in detail in Chapter 2, and effects of additional remedial physical education are analyzed in Chapter 3. Chapter 4 through Chapter 6 further explore longitudinal data to elucidate interrelation between physical fitness and anthropometrict measures (see Chapter 4), growth (see Chapter 5), and executive function (see Chapter 6). Specific details for exploratory analyses are described in Section 2.8.

2 Study design and methodology

This chapter provides information on the study design and methodology as well as selected group and participant characteristics.

2.1 Main focus of the SMaRTER study

The framework for the implementation of remedial physical education was provided by the resolution of the German Conference of Ministers of Education and Cultural Affairs on principles for the implementation of remedial physical education and for the qualification of remedial physical education teachers. According to these principles, remedial physical education should (1) focus on pupils with motor and psychosocial deficits, (2) be taught by specially trained physical education teachers, (3) aim to restore the physical performance of participating children to a level comparable to that of their peers, (4) include children following specific pedagogical criteria for sports development, (5) aim to increase overall academic learning and performance, and (6) facilitate integration of students into school life (KMK, 1999). In line with these principles, the University of Potsdam, in cooperation with the Ministry of Education, Youth and Sports of the federal state of Brandenburg and the AOK Nordost, conducted the **SMaRTER** Study ("Überprüfung der Effekte von **S**portförderunterricht auf die **m**otorische und kogni**t**ive **E**ntwicklung von Grundschulkinde**r**n im Land Brandenburg"). The aim of this study was twofold:

1. Development of a curriculum for remedial physical education for primary school children in the federal state of Brandenburg according to training science principles with a focus on third and fourth grade children with deficits in physical fitness.
2. Analysis of the short-, mid-, and long-term intervention effects of a remedial physical education intervention in third grade children with deficits in their physical fitness.

2.2 SMaRTER study design

The remedial physical education intervention of the SMaRTER study started in the second semester of third grade of the 2018/19 school year and was conducted using a cluster-randomised controlled design. Accordingly, participating schools were randomly assigned either the intervention first and then control condition (i.e., INT-CON) or the control condition first and then intervention (i.e., CON-INT). During both intervention periods, the assessments were administered before and after each intervention/control, that is, at the beginning and end of the second semester of grade three (i.e., t0: February 2019 & t1: June 2019) and at the beginning and end of the first semester of grade four (i.e., t2: August 2019 & t3: January 2020). Due to primary schools encompassing grades 1 to 6 in the state of Brandenburg, mid- and long-term intervention effects were examined at the beginning and end of the second semester of fourth grade (i.e., t4: February 2020 & t5: June 2020) and at the beginning and end of fifth grade (i.e., t6: August 2020 & t7: June 2021) and sixth grade (i.e., t8: August 2021 & t9: June 2022). The study design is shown in Figure 2.1.

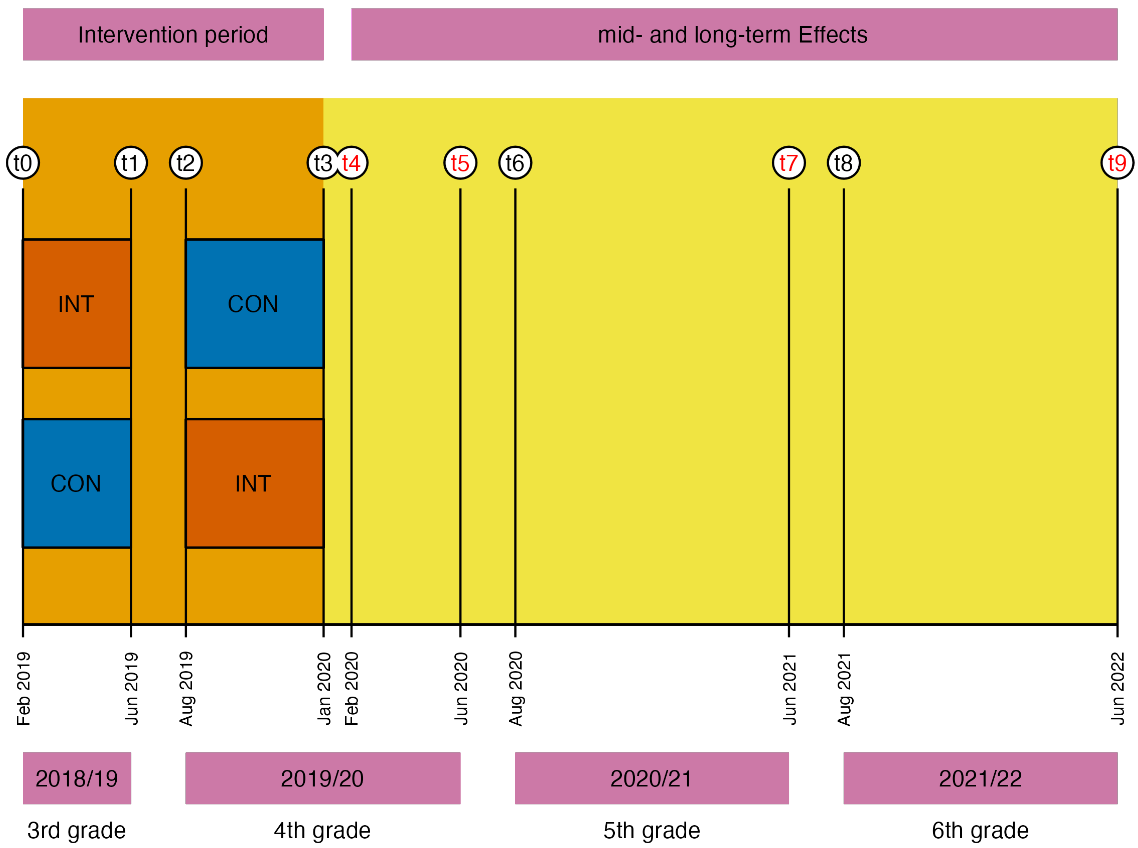


Figure 2. 1: SMaRTER study design; INT = intervention; CON = control

Due to the Covid 19 pandemic, however, the originally planned measurements of the mid- and long-term effects could not be realised. Measures to reduce infections (e.g., school closures, classes with/without restricted attendance and regulation of external visitors to the school, see Section 1.5) did not allow assessments t4, t5, and t7 (indicated in red in Figure 2.1). The examinations during the Covid 19 pandemic were carried out in direct consultation with the Ministry of Education, Youth and Sports Brandenburg, taking into account the additional workload of the teachers and schools involved and in compliance with all regulations of the SARS-CoV-2 Containment Regulations at the respective time of examination (e.g., rigorous SARS-CoV-2-testing of assessors, 2 m distance, and obligation to wear a mask). Furthermore, due to the changed framework conditions, a renewed consent of the legal guardians was obtained.

2.2.1 Remedial physical education intervention

The developed and implemented remedial physical education intervention can be found elsewhere[82]. It included a comprehensive set of exercises to improve physical fitness during a 14-week period with 45 minutes of remedial physical education twice a week. Each session included a warm-up, a main part and a cool-down with different emphases. In each session, all exercises were selected and organised with the main objective of promoting physical activity and fitness, embedded in a comprehensive pedagogical context that focused on interactions between children and reflection on their actions and accomplishments. In addition, each session was supplemented by homework assignments created in cooperation with "Henrietta's Moving School"[83] program of the AOK Nordost. Homework included exercises covering coordination, muscular strength, and muscular endurance. Table 2.1 shows the focus of the 2 x 45 minutes of remedial physical education per week as well as supplementary homework assignments[84]. All sessions were conducted by certified remedial physical education teachers (MBJS, 2011).

Table 2. 1: Remedial physical education intervention curriculum

week	focus	homework
1	getting to know each other through group games	dancing hands (coordination) & clapping with push-ups (strength)
2	perception of space and reaction to acoustic signals	laying 8 (coordination) & one-leg stand (coordination)
3	balance exercises	tightrope walk backwards (coordination) & tightrope walk with raised knees (coordination)
4	fitness exercises	planche (strength) & lunges (strength)
5	running, jumping, and throwing exercises	mountain climbing (endurance) & ball throwing (coordination)
6	running, jumping, and throwing exercises	rope skipping (coordination) & side planche (coordination)
7	individual and partner coordination exercises	our-field jumping (coordination) & alternating jumps (coordination)
8	creative rhythm exercises	one-leg stand (coordination) & clapping solo (coordination)
9	ball games	ball transport (coordination) & ball transport on the back (coordination)
10	fitness exercises	side crunches (strength) & squats (strength)
11	wrestling & brawling exercises	boxing (endurance) & random previous homework exercise
12	fitness exercises	earlobes (coordination) & jumping jacks (endurance)
13	exercises games	mountain jumping (endurance) & bending and stretching (strength)
14	fitness exercises	squat jumps (strength) & least successful previous homework exercise

2.3 Inclusion criteria

Following the principles for remedial physical education (KMK, 1999), only schools that employ certified remedial physical education teachers could be included[85] (MBJS, 2011). These schools were screened for children with deficits in their physical fitness to include in the study. A physical fitness deficit was identified using data from the EMOTIKON study[86] collected in the 2018/19 school year.

Using the six physical fitness tests from the EMOTIKON study, a physical fitness deficit was defined as:

(1) Performance in the first performance quintile (lowest 20%) in four of the six tests; or

(2) a quintile average across the performance quintiles of all six tests of ≤ 1.5 (lowest 30%).

If the data for the children was incomplete or unavailable, their teachers were consulted and a mutual decision on inclusion into the study was made. In addition to the inclusion of children with physical fitness deficits and in accordance with the principles for curative physical education (KMK, 1999), schools were given the opportunity to include children with psychosocial deficits or overweight. Children with severe physical and psychosocial disorders (e.g., physical disability, autism spectrum, bipolar disorder) were not included in the study, but were welcome to participate in remedial physical education intervention. Random allocation into INT-CON and CON-INT conditions occurred at school level (i.e., cluster randomisation) and was concealed. Participating children and remedial physical education teachers were aware of their assignment to the INT-CON and CON-INT condition due to the nature of the intervention. Assessors were aware which schools were assigned to the INT-CON and CON-INT conditions due to the design and organisation of the SMaRTER study. Written consent to participate in the SMaRTER study was obtained twice from the children's legal guardians[87], in accordance with the current Declaration of Helsinki.

2.4 Included assessments

Data on anthropometric measures, physical fitness, executive function, physical activity, and socioemotional well-being were collected at each assessment. The socioeconomic data of the children's guardians were only measured at the first and last (i.e., t0 and t9) assessments.

2.4.1 Anthropometric measures

Anthropometric measurements included standing and sitting height as well as body composition parameters. Standing and sitting height were measured with a stadiometer (seca 213, seca Gmbh, Hamburg, Germany) according to a standardised protocol. Body mass and body composition parameters (e.g., body fat percentage, muscle mass, and lean mass) were determined using a bioimpedance analysis system (InBody 720, BioSpace, Seoul, Korea). Based on these assessments and the age of the participants, the maturity offset according to Mirwald et al. (2002) and Moore et al. (2015) as well as the body mass index (BMI) were calculated. Formulas are provided in the Appendix (see Appendix Section 8.1.2 and Section 8.1.3).

2.4.2 Physical fitness

Physical fitness was assessed with the EMOTIKON test battery consisting of standing long jump, 20 m sprint, 6 min run, star run, ball push test, and one leg balance. In addition, hand grip strength as well as balance and gait parameters were tested in combination with a dual task.

2.4.2.1 EMOTIKON test

Standing long jump

Standing long jump was conducted to assess the muscular power of the lower limbs. Children had to jump as far as possible from the frontal stance and land with both feet together without any steps or arms touching the ground. Arm swings were allowed. The performance was scored using a tape measure as the distance in meters to the nearest centimeter. Two trials were performed, of which the better one was included in the analysis (Fühner et al., 2021, 2022; Golle, 2015; Golle et al., 2015). The standing long jump shows reliable intraclass correlation coefficients in children aged 5-12 years (ICC = .88 [95% CI .84 - .91] to .94 [95% CI .93 - .95]) (Fernandez-Santos et al., 2015; Fjørtoft et al., 2011) and is strongly associated with other jump tests such as the squat jump, the countermovement jump, and the Abalakov jump test (r = .73 - .78) (Fernandez-Santos et al., 2015).

20 m sprint

To assess linear sprint speed, a 20 m sprint was performed. The children had to sprint a distance of 20 m as fast as possible from an upright position in response to an acoustic signal. The performance was assessed as time to completion, accurate to 1/10 of a second, and was measured with a stopwatch. The best time from two trials was used in this analysis (Fühner et al., 2021, 2022; Golle, 2015). Test-retest analyses show high reliability for the 20m sprint in children aged 6 - 18 years (r = .71 - .9) (Bös et al., 2009; Fjørtoft et al., 2011).

6 min run

A 6 min run was performed to assess cardiorespiratory fitness. The test was conducted on a 54 m track (around a volleyball court: 18 m * 9 m) and children had to run as far as possible at a self-determined speed within 6 minutes. Performance was assessed as the maximum distance covered to the nearest 9 m (Fühner et al., 2021, 2022; Golle, 2015). This test shows high test-retest reliability in children aged 5 to 18 years (r = .72 - . 92) (Bös et al., 2009; Fjørtoft et al., 2011; Lawrenz & Stemper, 2012) and correlates moderately with other established proxies of cardiorespiratory fitness such as VO2max[88] (r = .46 - .69) and the shuttle run[89] (r = .74 - .83) (Faude et al., 2004; Haaren et al., 2011; Lawrenz & Stemper, 2012).

Star run

The star run assessed children’s coordination under time pressure. In this test, children had to run from a central position in a star-shaped area on a 9 x 9 m field, using different running styles (i.e., forward, backward, and lateral steps) according to a given protocol (see Figure 2.2). The center and the different spikes of the star were marked by pylons which had to be touched with the hand. The total distance to be covered was 50.912 m and had to be completed as quickly as possible. The performance was evaluated as the fastest time to the nearest 1/10 s from two test trials (Fühner et al., 2021, 2022; Golle, 2015; Golle et al., 2015). The star run is reliable (test-retest) in children aged 8 - 10 years (ICC = .68 [95% CI .53 - .79]) (Schulz, 2013).

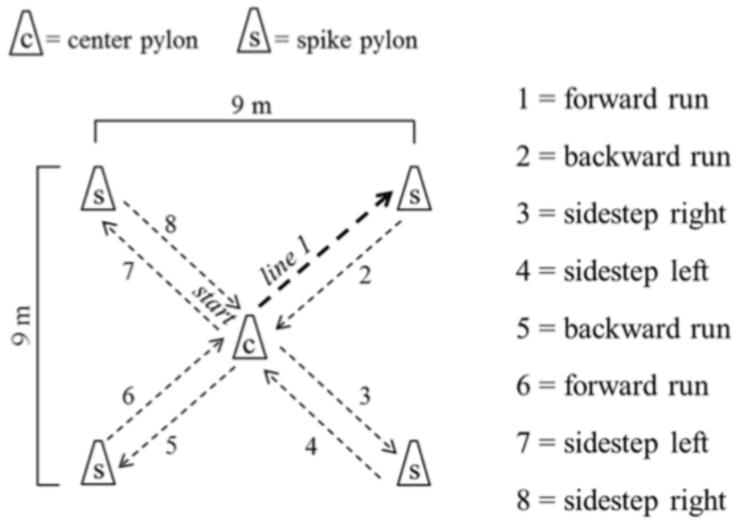


Figure 2. 2: Star run set up and protocol (adapted from Golle, 2015, p. 61)

Ball push test

A ball push test was performed to assess the muscular power of the upper limbs. Children had to push a 1 kg medicine ball in front of their chest with both hands from a standing position as far as possible. The performance was evaluated as the best of two maximum ball-pushing distances to the nearest ten centimeters (Fühner et al., 2021, 2022; Golle, 2015). Test-retest analysis showed the ball push test to be reliable in children aged 8 to 10 years (ICC = .81 [95% CI .71 - .87]) (Schulz, 2013).

One leg balance

The one leg balance test assessed static balance. Children were instructed to stand on their preferred leg for 60 seconds with their hands on their hips, eyes closed, and the other leg lifted forward at a hip angle of 60-90°. This position had to be held after an acoustic signal without releasing the hands from the hips, touching the lifted leg with the supporting leg or the floor, or bouncing/tapping the supporting leg. Performance was assessed by the time spent in the position to the nearest 1/10th of a second. If the child achieved a time < 5 seconds, a second attempt was allowed (Bormann, 2016; Granacher & Golle, 2016). Test-retest analyses showed that this test is reliable in children aged 7 - 10 years (ICC = .69 [95% CI .61 - .75]) (Bormann, 2016).

2.4.2.2 Additional fitness tests

Hand grip strength

Isometric handgrip strength was used to assess upper limb muscle strength and measured using a digital hand dynamometer (JAMAR® Plus+, Sammons Preston, Bolingbrook, IL, USA) on the dominant hand only. Children were in a seated position with the elbow of the dominant hand flexed at 90° to the side of the body and were instructed to squeeze as hard as they could for three seconds (Wind et al., 2010). High test-retest reliability was found in children and adolescents (ICC = .94 and .98, respectively) (Gerodimos, 2012).

Gait test

The gait test assessed postural balance in combination with a cognitive interference task. Gait parameters were measured on a 10 m instrumented walkway equipped with an OptoGait optoelectronic system (Micrograit, Bolzano, Italy). In order to avoid speed changes at beginning and end of the walkway, two additional meters were added at the beginning and end of the walkway. In total, one test run and two trial runs were performed for both the single task (gait only) and the double task (gait and inference task). The gait parameters for each condition were averaged over the two test runs. The inference task consisted of counting backwards from a random number between 60 and 100 in steps of three. A new number was given for each test run and the number of calculations and errors were documented (Beurskens et al., 2015). If the child could not perform the calculation in this test, successively simplified versions were implemented until the child could perform the calculations. Simplifications consisted of lowering the starting number to a random number between 10 and 20 and if necessary further counting backwards in steps of one.

Balance test

The balance test assessed static balance in combination with a cognitive inference task. The balance test was performed on a force plate (Leonardo, Kistler, AMTI) to determine parameters of the center of pressure. Children had to balance on one leg with eyes open, with the raised leg bent at the knee at ~ 60° for 30 s (Figura et al., 1991). Similar to the EMOTIKON one leg balance test, the criteria for a failed attempt were release of the hands from the hips, contact of the raised leg with the supporting leg or the floor, and bouncing/tapping of the supporting leg. The same arithmetic interference task as for the gait test was utilised and one test and one trial run each with single and dual task conditions were performed.

2.4.3 Executive function

The children's executive functions were assessed using the digit symbol substitution test, the trail making test, and the Simon task.

Digit symbol substitution test

The digit symbol substitution test was administered to assess attention and psychomotor processing speed. The children were presented with numbers from 1 to 9, with a symbol assigned to each number (see Figure 2.3). During the test, the children had 90 seconds to correctly assign the symbols to as many consecutive numbers as possible (Petermann & Petermann, 2011). The test was administered in a pen and paper version and performance was scored as the number of symbols correctly assigned in 90 seconds.

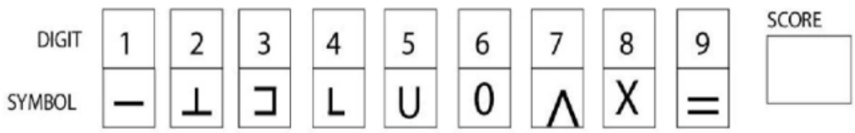


Figure 2. 3: Substituted symbols to the numbers 1 to 9 in the digit symbol substitution test (adapted from Health ABC, 2005, p. 9)

Trail making test

The trail making test was used to assess mental flexibility and fine motor abilities, and consisted of two parts (i.e., version A and version B). In version A, children had to connect the numbers from 1 to 15 in ascending order as quickly as possible and without lifting the pencil. In version B, they had to connect the numbers from 1 to 8 and the letters from A to G in alternating ascending order, starting from 1, as quickly as possible without lifting the pencil

(R. Reitan, 2004; Reitan, 1971; Reitan & Wolfson, 1995). Both versions were implemented in pen and paper and performance was evaluated as time to completion and number of errors.

Simon task

The Simon task was implemented to assess inhibitory control of executive functions (Simon & Rudell, 1967). The implemented version of the Simon task was modeled after the "Simon paradigm" by von Bastian et al. (2016). Their code[90] is available online and served as the source, which was converted into an app for the IPad. In this test, children were presented with stimuli (i.e., coloured circles) that varied according to their position on the screen (i.e., left or right) and colour (i.e., red or blue). The colour of the stimuli had to be identified as quickly as possible using fixed buttons at the bottom of the screen. Accordingly, the correct response button and the stimulus position could be congruent (i.e., stimulus on the same side as the correct answer button) or incongruent (i.e., stimulus on opposite side as the correct answer button; see examples in Figure 2.4). Six test runs were conducted, followed by three blocks of 20 stimuli as trial runs. Of note, to ensure a balanced test design, 30 test runs each were coded as congruent and incongruent and randomised in order to create a fixed protocol. This protocol was implemented on each trial, meaning that each child performed the same order of trials on each trial. Since the number of stimuli for side and colour was not set at 30 each, their distribution was slightly skewed towards blue stimuli (i.e., 32 blue and 28 red) and more severely skewed towards stimuli on the left side (i.e., 40 left and 20 right). Performance was evaluated as the mean reaction time for each trial averaged for the congruent and incongruent condition. The test was conducted with an IPad 6 (OSX 15.2.1, Apple Inc., Cupertino (CA), USA). The Simon effect can already be observed in children as young as 4 years old (Davidson et al., 2006) and is a reliable measure of inhibition in young adults (split-half reliability == .8) (Bastian et al., 2016).

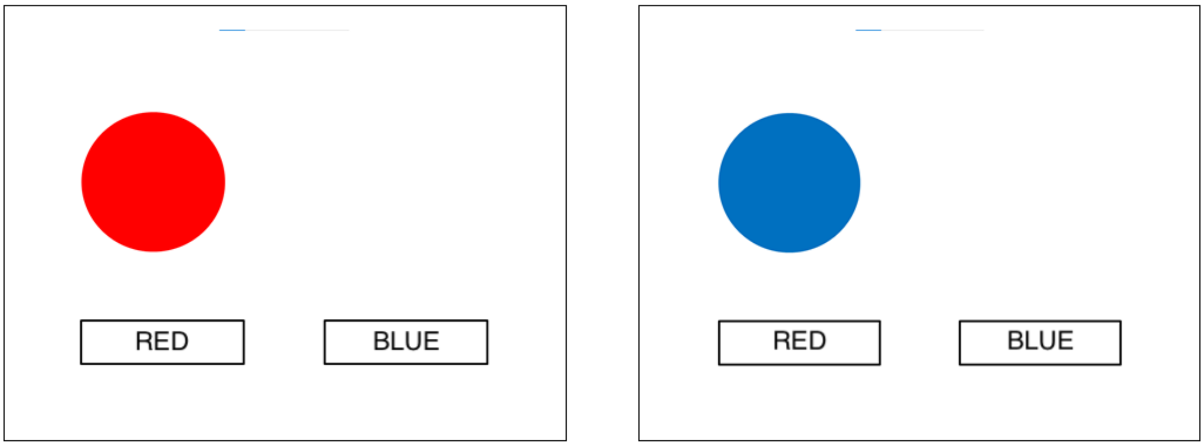


Figure 2. 4: Examplatory Simon task trials in congruent (left) and incongruent (right) condition

2.4.4 Physical activity

Physical activity was recorded with a pedometer and a physical activity questionnaire. The pedometer (Speedy, Kasper & Richter GmbH, Uttenreuth, Germany) utilised a three-dimensional motion sensor to count the number of daily steps. Pedometers were worn for seven days after each examination and children were instructed to wear them at all times. Their displays were covered to avoid interference between the number of steps displayed and the children's behaviour. Pedometers have been shown to be a reliable assessment of physical fitness with high compliance in children (Clemes & Biddle, 2013). The physical activity questionnaire used was the “MoMo Activity Questionnaire” for children and adolescents. The questionnaire focuses on time spent in organised physical activity in schools, clubs, and outside clubs, as well as in daily life, and converts this information into gaily minutes spent on MVPA. It is a reliable tool for assessing physical activity in children (Bös et al., 2009; Jekauc et al., 2013; S. Schmidt et al., 2016).

2.4.5 Socioemotional well-being

Socioemotional well-being was assessed using the KidKindl questionnaire for children aged 7 to 13 (Ravens-Sieberer et al., 2007). The questionnaire assesses the child’s subjective well-being in relation to body, psyche, self-esteem, family, friends, and school with four questions each and answers on a scale from 0 (never) to 4 (always). The questionnaire has proven to be a reliable assessment of socioemotional well-being, with slightly higher scores for the version with a legal guardian (Ellert et al., 2011; Erhart et al., 2009). The questionnaire was completed by the children themselves and one of their legal guardians.

2.4.6 Socioeconomic status

Socioeconomic status was determined using the KiGGS questionnaire, which focused on the guardians’ schooling, professional qualifications, occupational status, employment, and income (Lampert et al., 2014). The questionnaire was filled out by all available guardians and answering the questionnaire was voluntary.

2.5 Descriptive group and study characteristics

A total of 76 3rd grade students from eleven different primary schools in 9 different districts of the federal state of Brandenburg participated in the SMaRTER study. Six schools (44 children; 18 girls and 26 boys) were randomly assigned to the INT-CON group and 5 schools (32 children; 17 girls and 15 boys) to the CON-INT group. The number of children enrolled in each school ranged from 4 to 11, with a median of 6. The mean age at baseline was 9.2 ± 0.5 years and mean maturity offset was -2.9 ± 0.8 years according to Mirwald et al. (2002) and -2.9 ± 0.6 years according to Moore et al. (2015). The group- and gender-specific baseline-characteristics are reported in Table 2.2 and depicted in Figure 2.5.

Table 2. 2: Baseline characteristics presented for group and gender

	INT-CON		CON-INT	
	girls	boys	girls	boys
n	18	26	17	15
age (years)	9.1±0.7	9.3±0.5	9±0.4	9.1±0.5
maturity offset (years)	-2.5±0.6	-3.2±0.5	-2.2±0.8	-3.5±0.5
body mass (kg)	30.1±7.8	41.1±13.4	36.2±12.6	35.9±11.9
body height (cm)	133.9±6.5	141.3±6.1	139±9.9	139.3±7.5
BMI (kg/m²)	16.6±3	20.4±5.5	18.3±4.1	18.3±5.3

Participants baseline characteristics; n = number of included children; INT = intervention; CON = control condition; BMI = body mass index

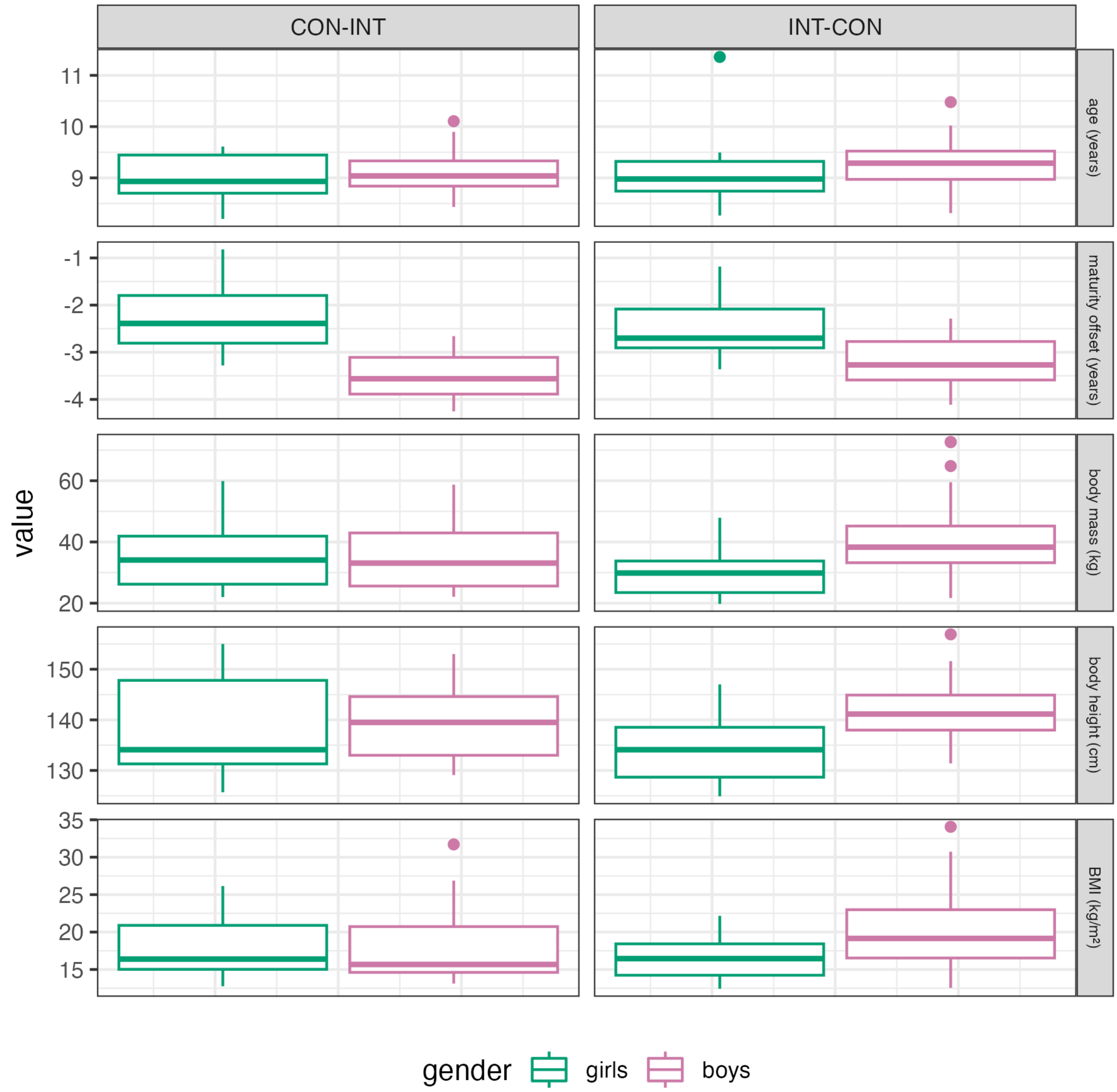


Figure 2. 5: Boxplots of anthropometric baseline characteristics depicted for group and gender; CON = control condition; INT = intervention; BMI = body mass index

2.5.1 Maturity offset

Although the participating children had similar ages at baseline, the maturity offset differed especially between girls and boys. Centering the baseline age and maturity offset at 0 according to the group mean showed almost all girls to have a higher maturity offset compared to their age, while the boys' maturity offset was lower, with only minor differences between the different formulas. This is illustrated in Figure 2.6.

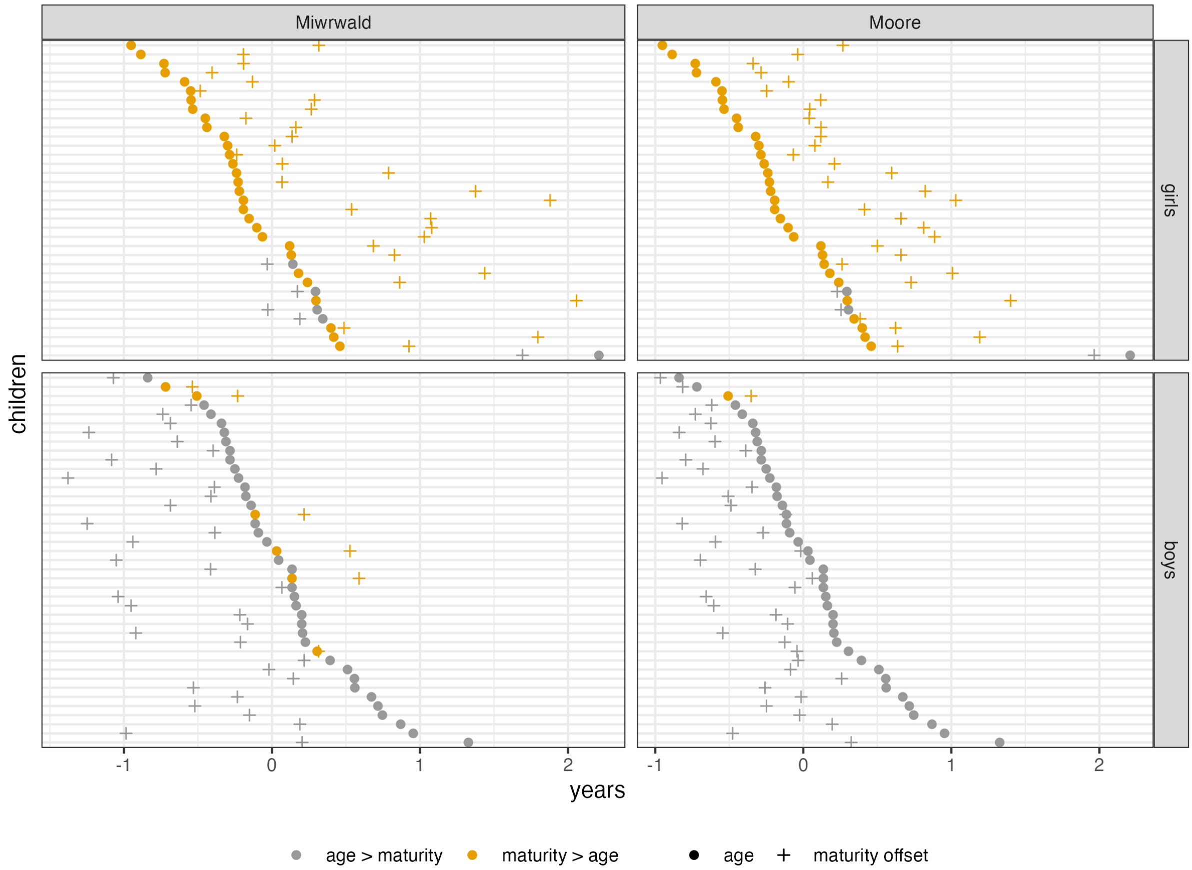


Figure 2. 6: Difference between age and maturity offsets according to Mirwald (Mirwald et al., 2002) and Moore (Moore et al., 2015) in boys and girls, both centered at 0

2.5.2 Socioeconomic information

The socioeconomic data of the guardians were only collected at the beginning of the study, as the last assessment could not be completed (see Figure 2.1). 117 guardians provided a total of socioeconomic information corresponding to 68 children. For 19 children, one guardian's information and for 49 children, two guardians' information was provided. As socioeconomic data was not consistently available for all children and is likely to be biased is it was requested and provided on a voluntary basis (Goyder et al., 2002), socioeconomic information is only reported as a sample descriptor but not further included in the analyses.

61 guardians were female, 54 male (2 missing), the average age was 40 ± 7 years (10 missings), 112 reported being German, and one each was Italian, Afghan, and Serbian (2 missings). In 17 guardians the highest school degree was "Hauptschulabschluss", in 31 "Realschulabschluss", in eight "Abschluss der polytechnischen Oberschule", in 17 "Fachhochschulreife", in 34 "Hochschulreife/Abitur", in one "erweiterter Hauptschulabschluss", and two did not graduate from any school (8 missings). 53 guardians did full-time paid labour (42 ± 8.6 hours/week), 36 did part-time paid labour (29 ± 6 hours/week), and 12 were unemployed. The remaining guardians were in training (1), in a mini-job or occasionally employed (5), or on parental leave (1) (8 missings). Regarding working hours per week, one guardian reported 168 working hours/week, noting that their unpaid care work was equivalent to a 24/7 working week. The information on income was provided by the legal guardians of 46 participating children. The average household income[91] was 3,338.84 ± 1,444.97 € and was earned on average by two individuals per household (range 1 to 4) with an average income[92] of 1,819.49 ± 677.78 € (2

missings). This resulted in an average household income per capita of 855.31 ± 438.68 € in the study sample. The distribution of household income and household income per capita within the study sample is shown using a Lorenz curve[93] in Figure 2.7. Furthermore, the Gini index was calculated for the sample according to the formula of the World Bank (Bank, 2023a). The Gini index provides information on the equality of income distribution within a group[94] and was calculated for household income (Gini index = 24.32) and household income per capita (Gini index = 27.51). Using country-specific reference values, Gini indices of 50 to 70 indicate very unequal income distribution, while countries with a Gini index of 20 to 35 indicate relatively equal distribution (Willis, 2020).

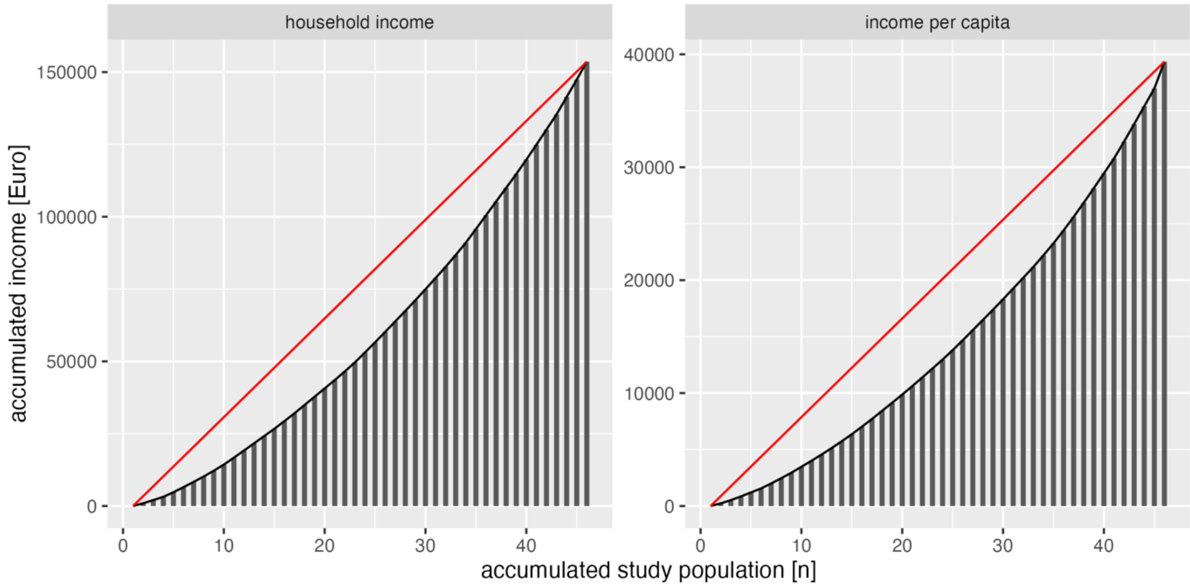


Figure 2. 7: Distribution of household income and income per capita illustrated using Lorenz curve plots

The Gini indices and Lorenz curves of the sample indicate that household income and income per capita are relatively evenly distributed within the sample. However, in 2019, the average annual household income in the federal state of Brandenburg was €60,366 and the average income per capita was €23,984 (Amt für Statistik, 2021). This results in an average monthly household income of €5,030.50 and an income per capita of €1,998.67[95]. Compared to the included sample, only 6 of the 46 households had above average household income and only one child had values above the average income per capita. To illustrate the discrepancy, a Lorenz curve was plotted with an adjusted 45° line for the average income in Figure 2.8.

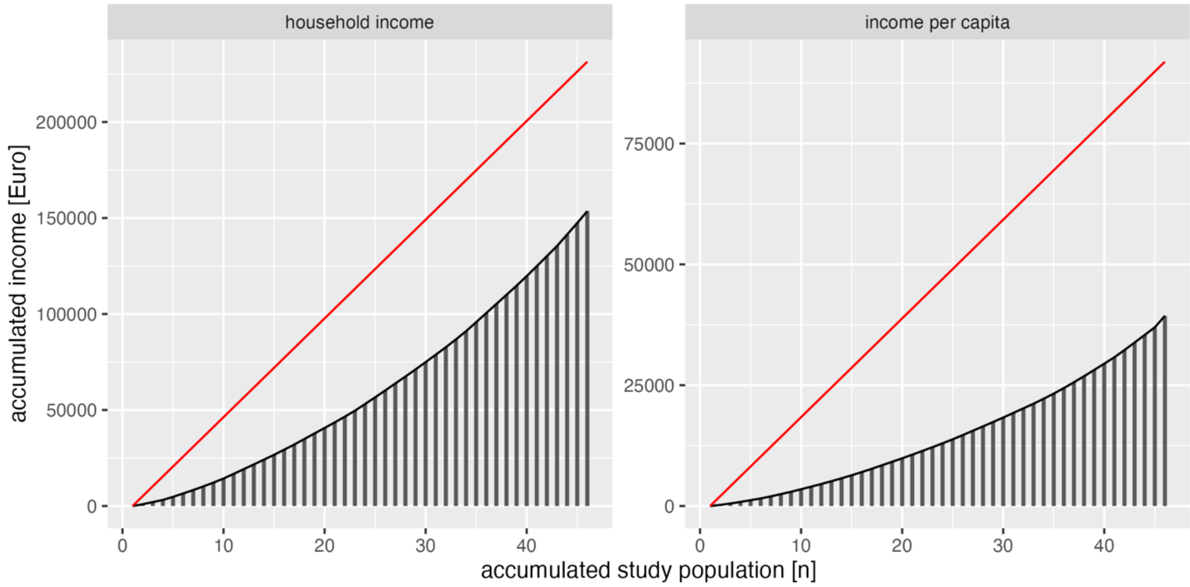


Figure 2. 8: Distribution of household income and income per capita illustrated using Lorenz curve plots with adjusted 45° line for the average household income income and income per capita of the federal state of Brandenburg in 2019

2.6 Study dropout

Continuous drop-out from the SMaRTER study is shown in Figure 2.9. Reasons given for dropping out between t0 and t3 were changing schools (4 children) and the increased time required for the additional remedial physical education classes of the SMaRTER study (6 children). Students who dropped out between t6 (16 students) and t8 (13 students) withdrew (or no longer gave consent to participate in the study) due to additional workload and/or increased risk of SARS-CoV-2 infections due to the Covid 19 pandemic.

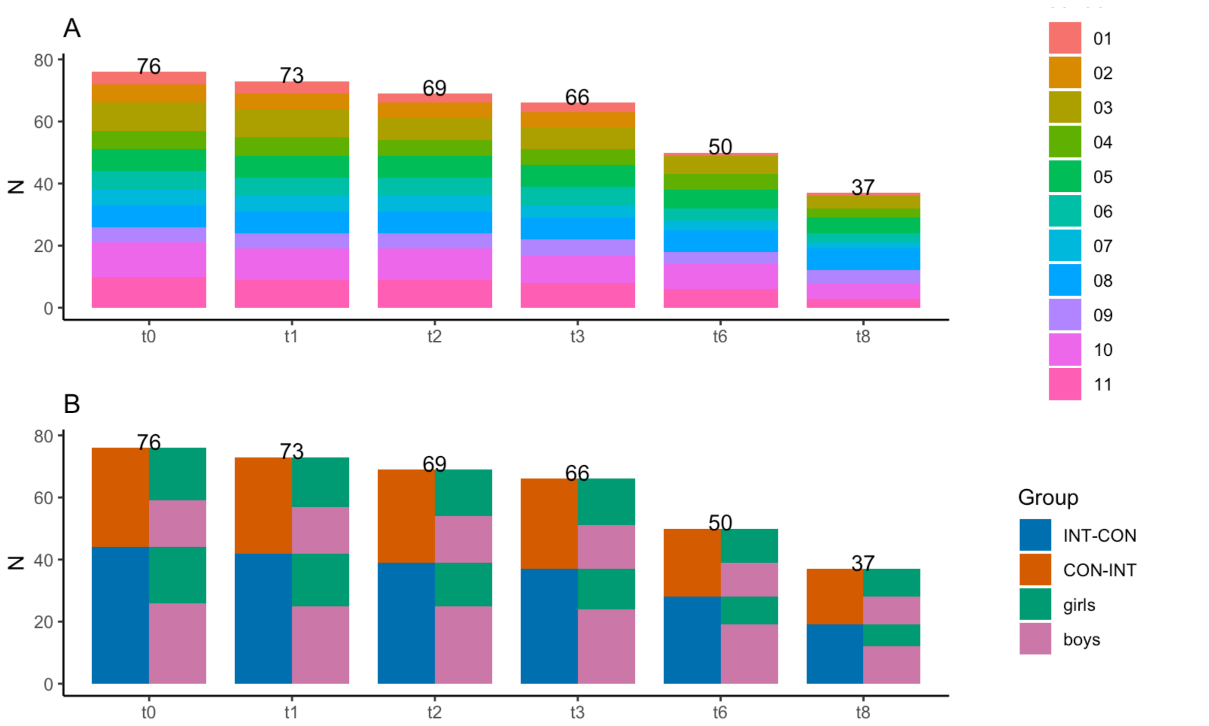


Figure 2. 9: Depictions of dropouts across all assessments A Dropout in relation to schools; B Dropout in relation gender per intervention group; N = number of participants left in sample; INT-CON = intervention control; CON-INT = control intervention

2.7 Anthropometric group differences

Before analysing intervention effects in physical fitness or executive function, possible group and gender differences were analysed for the variables age, maturity offset (according to Mirwald et al. (2002)) and the anthropometric measures body height, body mass, and BMI (presented in Figure 2.10).

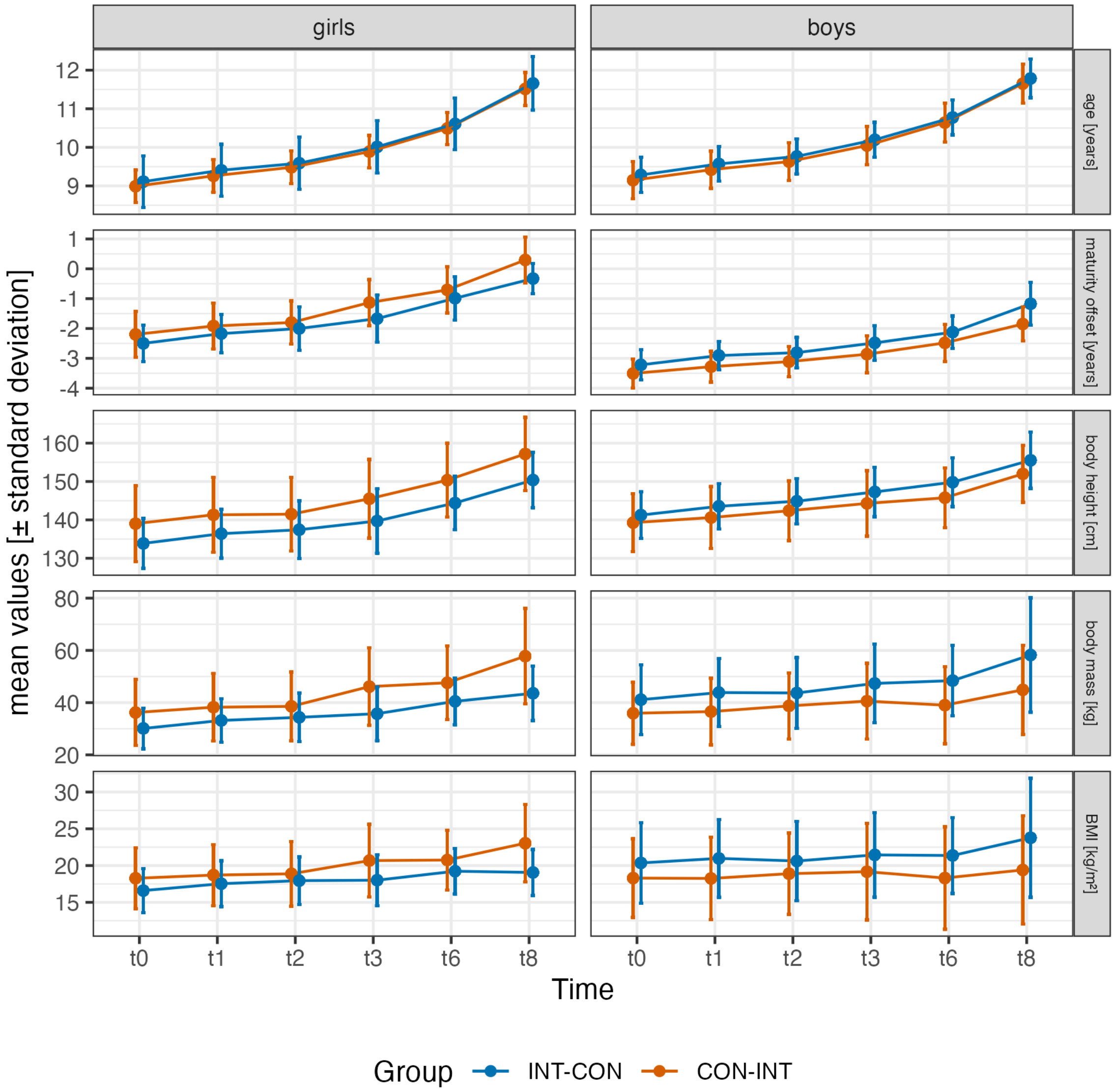


Figure 2. 10: Group and gender differences over time for age, maturity offset (Mirwald), body height, body mass, and body mass index (BMI)

Since the different displayed values are highly interdependent, meaning partly computed from each other, the differences between the genders and groups were investigated for the BMI only using LMMs.

The LMMs were adjusted with BMI as the dependent variable and the interaction of Group x Gender x Time as fixed factors. Child and School were included as random factors. Successive difference contrasts were set for the variables Group, Gender, and Time.

LMM revealed that (1) the BMI of boys in the INT-CON group is higher compared to boys in the CONINT group, while the reverse is true for girls, and (2) the BMI of girls in the CON-INT group shows a greater improvement between t2 and t3 and between t6 and t8 compared to girls in the INT-CON group, while the reverse is true for boys between t6 and t8. The results are shown in Table 2.3 and depicted in Figure 2.10. Accordingly, the BMI is included in the models to evaluate the intervention effects. Missing values are replaced by earlier available values. Due to dropouts, the Group x Gender x Time interactions for BMI were additionally tested using LMMs for all children who were present at t1 and t3. Results of these LMMs are reported in the Appendix Section 8.2.1 and show the same pattern.

Table 2. 3: Group x gender x Time interactions for body mass index across all assessments

Predictors	Est.	SE	z	p
Grand mean	19.26	0.83	23.30	**<0.001**
Group2-1	0.35	1.65	0.21	0.832
gender2-1	-2.12	1.05	-2.01	**0.045**
Time2-1	0.17	0.14	1.17	0.243
Time3-2	0.31	0.15	2.09	**0.038**
Time4-3	0.43	0.16	2.71	**0.007**
Time5-4	0.53	0.18	2.91	**0.004**
Time6-5	0.86	0.20	4.33	**<0.001**
Group2-1 * gender2-1	-4.77	2.11	-2.26	**0.024**
Group2-1 * Time2-1	0.03	0.29	0.09	0.924
Group2-1 * Time3-2	-0.44	0.30	-1.47	0.143
Group2-1 * Time4-3	-0.17	0.32	-0.54	0.588
Group2-1 * Time5-4	0.47	0.36	1.30	0.196
Group2-1 * Time6-5	-0.45	0.40	-1.15	0.251
gender2-1 * Time2-1	0.24	0.29	0.83	0.409
gender2-1 * Time3-2	0.08	0.30	0.26	0.791
gender2-1 * Time4-3	0.17	0.32	0.54	0.591
gender2-1 * Time5-4	0.03	0.36	0.07	0.941
gender2-1 * Time6-5	-0.04	0.40	-0.10	0.924
Group2-1 * gender2-1 * Time2-1	0.10	0.58	0.18	0.857
Group2-1 * gender2-1 * Time3-2	0.16	0.60	0.27	0.786
Group2-1 * gender2-1 * Time4-3	-1.49	0.64	-2.33	**0.021**
Group2-1 * gender2-1 * Time5-4	0.57	0.73	0.79	0.429
Group2-1 * gender2-1 * Time6-5	-2.31	0.79	-2.91	**0.004**
Random Effects				
σ2	0.70			
τ00Child	17.68			
τ00school	4.63			
ICC	0.97			
N Child	76			
N school	11			
Observations	348			
Marginal R2 / Conditional R2	0.133 / 0.974			
Est. = estimate; SE = standar error				

2.8 Analysis plan

The first analysis focuses on the effects of remedial physical education on physical fitness and executive function and is reported in Chapter 3.

Although various assessments of physical fitness were available for the analysis, this thesis focuses on the EMOTIKON tests as the main outcome. Focusing on EMOTIKON tests allowed for a standardisation of data using the population mean and standard deviation previously published (Fühner et al., 2021, 2022). In addition, based on the findings of Fühner et al. (Fühner et al., 2021, 2022), physical fitness is understood as a latent construct, represented by the four EMOTIKON tests: standing long jump, 20 m sprint, 6 min run and star run. Additionally, executive function as a secondary outcome is included in the analysis, consisting of the trail making test (i.e., time to completion for version A and B) (R. Reitan, 2004; Reitan, 1971; Reitan & Wolfson, 1995), the digit symbol substitution test (Petermann & Petermann, 2011), and the Simon task (i.e., mean reaction time in the congruent and incongruent condition) (Simon & Rudell, 1967). Of note, other physical fitness assessments (i.e., hand grip strength, gait test, and balance test), as well as physical activity and socio-emotional well-being were assessed for potential intervention related effects, but are not included in this report. Significant intervention-related effects will be included in the following analyses.

The second analysis explores the relationship between anthropometric parameters and physical fitness and is conducted in Chapter 4. Due to internal calculations, the anthropometric parameters estimated with InBody showed unexpectedly high correlations between body mass and other measures such as fat mass ($r = .95$), muscle mass ($r = .92$) and the sum value of muscle and fat mass ($r = 1$; see Appendix Section 8.4.1), indicating that the values are estimated applying a linear transformation model undisclosed by the manufacturer. Therefore, only body height and mass were analysed in relation to physical fitness. In addition to the physical fitness tests included in the first analysis, this analysis further included ball push test. This test was included based on the assumption that a higher body mass might be beneficial for test performances, thus revealing a potential positive relationship compared to the other tests, where a higher body mass is expected to be detrimental to test performances (see Section 1.2 for details).

The third analysis is focused on the relationship between maturity offset and physical fitness and is conducted in Chapter 5. The main goal is to analyse the effect of the maturity offset in relation to age on physical fitness. This is tested for the maturity offset according to Mirwald et al. (2002) as well as Moore et al. (2015). Physical fitness follows the same conceptualisation the intervention analysis in Chapter 3.

The fourth analysis examines the interaction between physical fitness and executive function (Chapter 6). Due to the availability of raw data from the Simon task, there was a unique opportunity to evaluate a high number of data points (i.e., 60 trials per child per evaluation), which are not available for digit symbol substitution and trail making test (i.e., 1 or 2 data points per child per evaluation). This allowed for a sensitive analysis of physical fitness parameters as variance components of an executive function task. Accordingly, executive function was defined as reaction times of each Simon task trial. Effects of physical fitness on executive function were assessed for all EMOTIKON tests to allow for a more distinct understanding of how different aspects of physical fitness might be associated with executive function.

2.9 Statistical information

Data preprocessing and analyses in Chapter 3, Chapter 4, and Chapter 5 was done with R [4.2.1; (R Core Team, 2022)]. Chapter 6 additionally used Julia [1.8.2, (Bezanson et al., 2017)].

Utilised package for preprocessing in R was the 'tidyverse' package (Wickham et al., 2019). Liner mixed models (LMM) estimation, supplementation, and post-processing in R was done using the 'lme4' package (D. Bates et al., 2015), the 'MASS' package (Venables & Ripley, 2002), the 'remef' package (Hohenstein & Kliegl, 2022), the 'sjPlot' package (Lüdecke, 2021), and the 'performance' package (Lüdecke et al., 2021). In Julia LMMs were estimated with the 'MixedModels.jl' package (D. Bates et al., 2023), and the 'MixedModelsExtras.jl' package (Alday, 2022) was used for data analysis and post-processing of LMMs. Plotting of data was done in R using the 'ggplot2' package (Wickham, 2016), the 'gridExtra' package (Auguie, 2017), and the 'car' package (Fox & Weisberg, 2019). Further, tabular presentation of data was done using the 'flextable' package (Gohel & Skintzos, 2022) and the 'officer' package (Gohel, 2022) in R.

Part II

Data-analyses and results

3 Analyses of intervention effects

3.1 Introduction

This chapters main concern is the analysis of the effects of the 14 week intervention of 2 x 45 minutes of remedial physical education implemented in the SMaRTER study described in Chapter 2. Improvements of physical fitness and executive function were expected for children who participated in the additional physical fitness compared to a passive control group.

3.2 Statistical analysis

The SMaRTER study implemented a cross-over design with control and intervention groups and pretest and multiple posttest assessments (see Section 2.2). The critical statistic concerns is a Group x Time interaction, with the expectation that children who participate in the additional remedial physical fitness interventions show larger physical fitness and executive function improvements from pre to posttest assessments compared to children of the control group.

The cross-over design was primarily adopted to also enable control children to benefit from possible intervention advantages in a time-delayed manner. The design also allowed for a second test of the primary research hypothesis. If there was an intervention benefit in the first period, it is expected that children in the control group will catch up. Of course, other outcomes are possible depending on whether the group doing the first intervention maintains or increases intervention benefits.

The study also allowed for an evaluation of long-term benefits of the intervention plan. There were additional tests at the beginning of fifth and sixth grade (see Figure 2.1). Due to dropouts, the number of children decreased from the first posttest to the second posttest to long-term follow-up tests (see Figure 2.9). To keep statistical power as large as possible, we tested these hypotheses with three separate linear mixed models (LMMs):

1. Intervention model 1 (INT M1): first intervention period only
2. Intervention model 2 (INT M2): pooled intervention periods for both groups[96]
3. Intervention model 3 (INT M3): middle and long term intervention-related effects

Specifically, we excluded dropout children for each model (i.e., INT M1 = t0 - t1, INT M2 = t0 - t3, INT M3 = t0 - t8; see Figure 3.1).

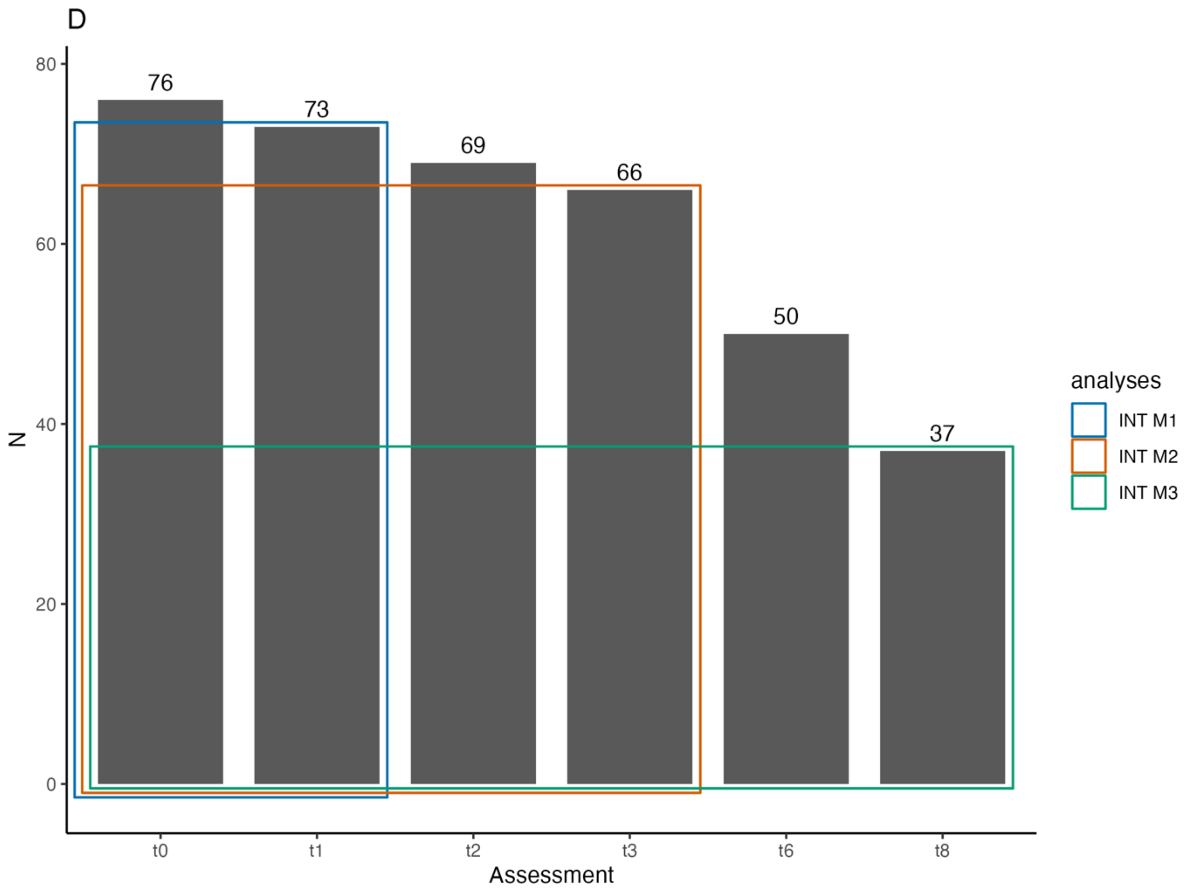


Figure 3. 1: Dropout with depiction of included children in different intervention models; N = number of assessed participants; INT M1 = Intervention model 1; INT M2 = Intervention model 2; INT M3 = Intervention model 3

For each of the three intervention models, main effects of Group and Time, as well as their interaction were examined with two separate LMMs, one for four tests of physical fitness (i.e., standing long jump, 20 m sprint, 6 min run, and star run) and one for five tests of executive function (i.e., trail making test version A, trail making test version B, digit symbol substitution test, Simon task congruent condition, & Simon task incongruent condition). Components of physical fitness and executive function included in both LMMS were z-transformed (see Appendix Section 8.1.4, and Section 8.1.5) and set as factor levels of Test in each LMM.

Age (centered at 9.3 years), gender, and body mass index (BMI; centered at 17.9 kg/m2 for girls and 19.8 kg/m2 for boys; see Section 2.7), as well as proximity to Berlin (R. Kliegl & Teich, 2022) were added as fixed effect control covariates. Significance of covariates was tested using log likelihood ratio tests (LRT) that compared LMMs with and without covariates. All covariates that did not significantly improve LMMs fit were removed from the LMM.

Child was specified as a random factor. Thus, LMMs take into account individual differences between children. Successive difference contrasts were set for the variables Test, Time, Group, Gender, and Berlin Proximity.

3.3 Intervention model 1: First intervention period

3.3.1 Physical fitness

3.3.1.1 Model fitting

Model selection is documented in the Appendix Section 8.3.1.1. For the analysis of intervention effects on physical fitness, Group and Time were included in the fixed factor structure nested within physical fitness Test. In addition, BMI was also included in the nested structure.

3.3.1.2 Results

The LMM revealed significant Time-related improvements for star run ($\beta = .31$; SE = .1; $z = 3.13$; $p = .002$). In addition, effects of BMI were significant for 6 min run ($\beta = -.09$; SE = .02; $z = -5.27$; $p < .001$), 20 m sprint ($\beta = -.06$; SE = .02; $z = -3.31$; $p = .001$), and standing long jump ($\beta = -.04$; SE = .02; $z = -2.09$; $p = .038$), with a lower BMI being associated with better performances. However, there was no evidence for interactions between Time and Group for any of the Tests. Results are presented in Table 3.1 and illustrated in Figure 3.2. Means and standard deviations are reported in the Appendix Section 8.3.1.2.

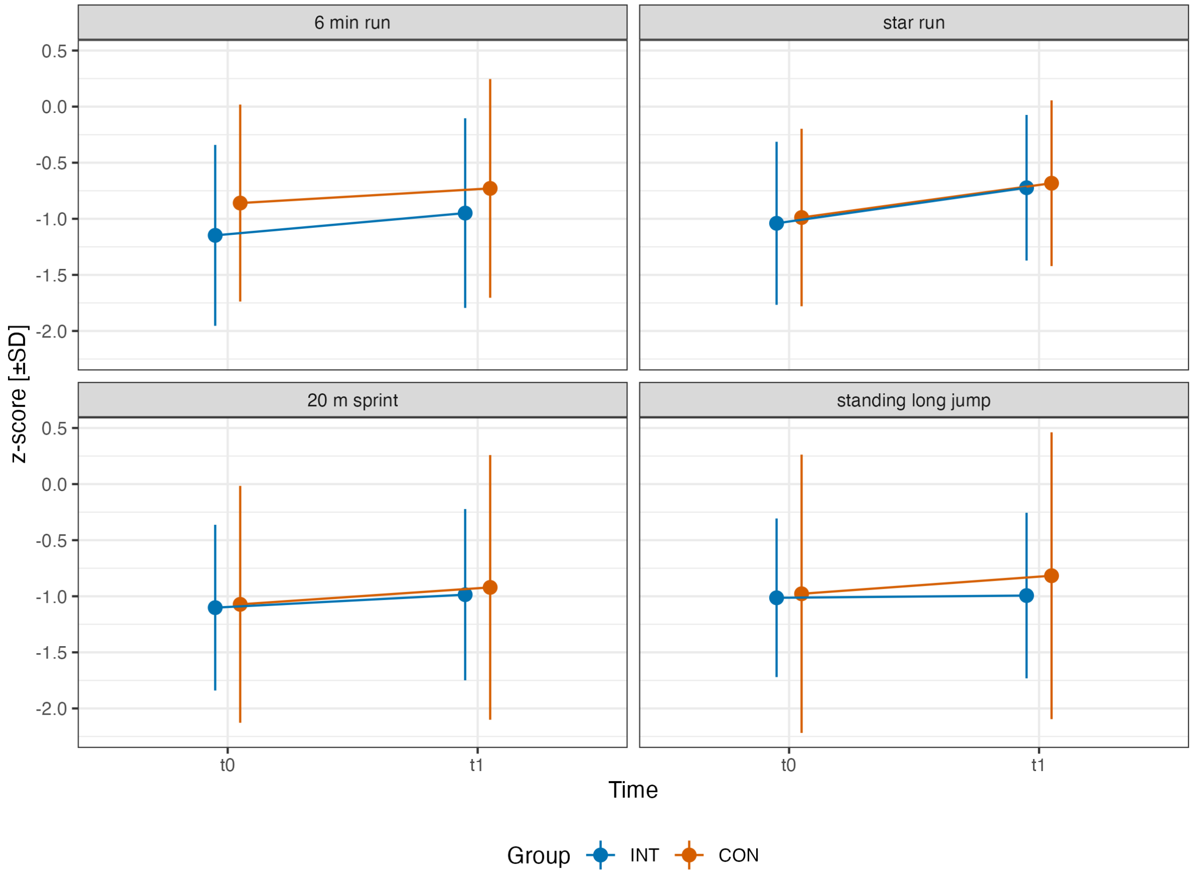


Figure 3. 2: Physical fitness plots for the first intervention period CON = control group, INT = intervention group.

Table 3. 1: Results of LMM for physical fitness for the first intervention period

Predictors	Est.	SE	z	p
Test [Run]	-0.93	0.09	-10.91	**<0.001**
Test [Star_r]	-0.86	0.09	-10.05	**<0.001**
Test [S20_r]	-1.02	0.09	-12.04	**<0.001**
Test [SLJ]	-0.95	0.09	-11.23	**<0.001**
Test [Run] : Group2-1	-0.21	0.17	-1.22	0.224
Test [Star_r] : Group2-1	-0.05	0.17	-0.30	0.761
Test [S20_r] : Group2-1	-0.02	0.17	-0.11	0.909
Test [SLJ] : Group2-1	-0.09	0.17	-0.55	0.580
Test [Run] : Time2-1	0.19	0.10	1.92	0.056
Test [Star_r] : Time2-1	0.31	0.10	3.15	**0.002**
Test [S20_r] : Time2-1	0.15	0.10	1.51	0.131
Test [SLJ] : Time2-1	0.10	0.10	0.99	0.323
Test [Run] : bmi_c	-0.09	0.02	-5.27	**<0.001**
Test [Star_r] : bmi_c	0.01	0.02	0.42	0.673
Test [S20_r] : bmi_c	-0.06	0.02	-3.31	**0.001**
Test [SLJ] : bmi_c	-0.04	0.02	-2.09	**0.038**
Test [Run] : Group2-1 * Time2-1	0.08	0.20	0.41	0.683
Test [Star_r] : Group2-1 * Time2-1	0.01	0.20	0.03	0.977
Test [S20_r] : Group2-1 * Time2-1	-0.03	0.20	-0.15	0.879
Test [SLJ] : Group2-1 * Time2-1	-0.15	0.20	-0.75	0.454
Random Effects				
σ2	0.34			
τ00Child	0.34			
ICC	0.50			
N Child	72			
Observations	561			
Marginal R2 / Conditional R2	0.123 / 0.559			

Run = 6 min run, Star_r = star run, S20_r = 20 m sprint, SLJ = standing long jump, bmi_c = gender-specific zero-centered body mass index, Est. = estimate; SE = standar error

3.3.2 Executive function

3.3.2.1 Model fitting

Model selection is documented in the Appendix Section 8.3.2.1. For the analysis of intervention effects on executive function, Group and Time were included in the structure of fixed factors nested in executive function Test. In addition, gender was included in the fixed effects structure.

3.3.2.2 Results

The LMM revealed a significant gender-related effect (β = .33; SE = .13; z = 2.53; p = .012), with better overall executive function performances in boys compared to girls. Overall, Time-related improvements were found for trail making test version A (β = .52; SE = .13; z = 4.05; p < .001), trail making test version B (β = .25; SE = .13; z = 1.97; p = .050), digit symbol substitution test (β = .36; SE = .13; z = 2.78; p = .006), and Simon task incongruent condition (β = .29; SE = .13; z = 2.26; p = .024). However, there was no evidence for interactions between Time and Group for any included executive function Tests. The results are presented in Table 3.2 and mapped in Figure 3.3. Means and standard deviations are reported in the Appendix Section 8.3.2.2.

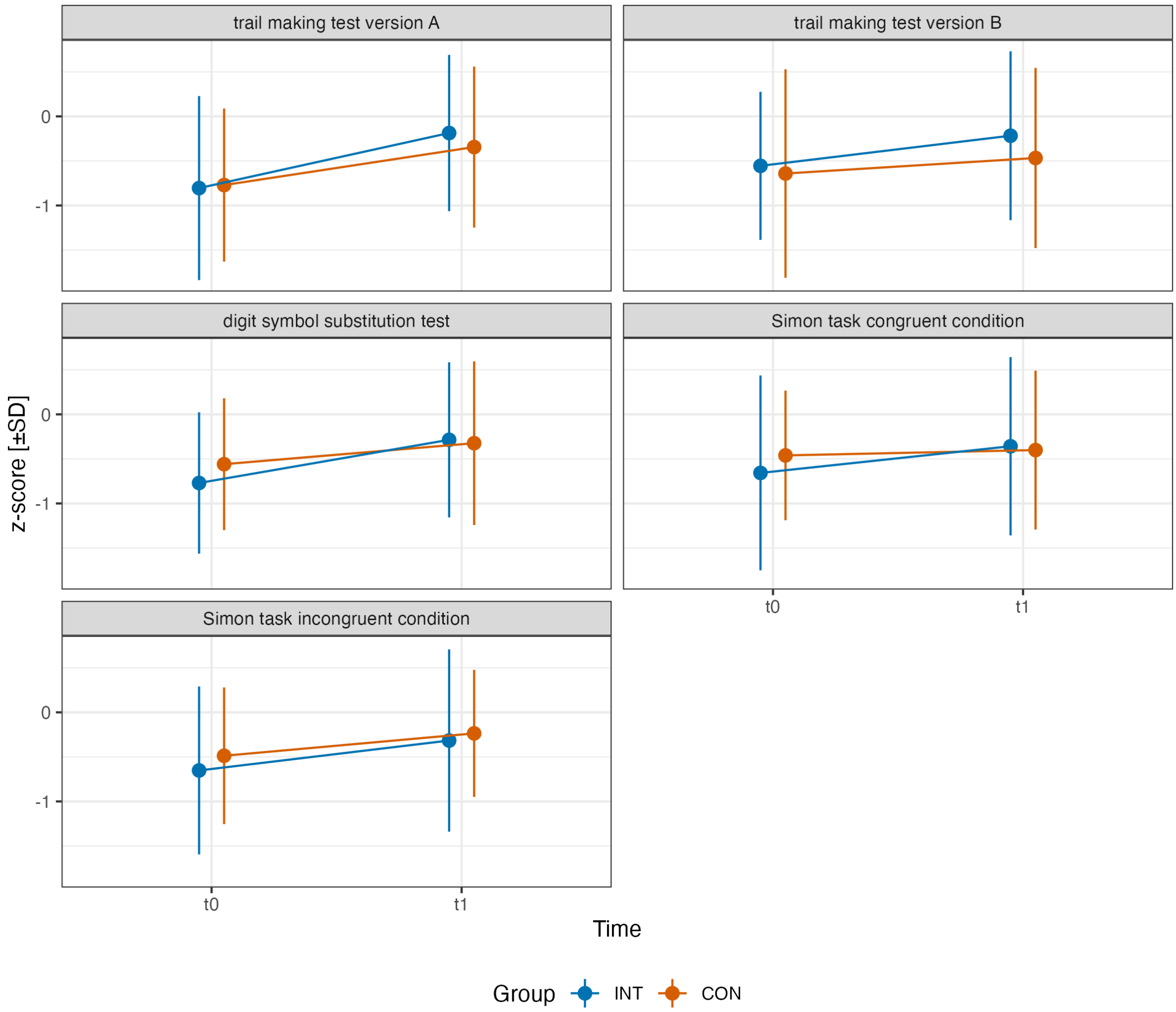


Figure 3. 3: Executive function performance plots for the first intervention period, CON = control group, INT = intervention group

Table 3. 2: Results of LMM for executive function performances for the first intervention period

Predictors	Est.	SE	z	p
Test [TMTa]	-0.54	0.09	-6.31	**<0.001**
Test [TMTb]	-0.49	0.09	-5.65	**<0.001**
Test [DSST]	-0.50	0.09	-5.82	**<0.001**
Test [SC]	-0.49	0.09	-5.64	**<0.001**
Test [SI]	-0.44	0.09	-5.11	**<0.001**
gender2-1	0.33	0.13	2.53	**0.012**
Test [TMTa] : Group2-1	-0.01	0.17	-0.05	0.962
Test [TMTb] : Group2-1	0.10	0.17	0.57	0.571
Test [DSST] : Group2-1	-0.16	0.17	-0.91	0.365
Test [SC] : Group2-1	-0.15	0.17	-0.85	0.396
Test [SI] : Group2-1	-0.19	0.17	-1.11	0.266
Test [TMTa] : Time2-1	0.52	0.13	4.05	**<0.001**
Test [TMTb] : Time2-1	0.25	0.13	1.97	0.050
Test [DSST] : Time2-1	0.36	0.13	2.78	**0.006**
Test [SC] : Time2-1	0.18	0.13	1.37	0.172
Test [SI] : Time2-1	0.29	0.13	2.26	**0.024**
Test [TMTa] : Group2-1 * Time2-1	0.13	0.26	0.52	0.602
Test [TMTb] : Group2-1 * Time2-1	0.11	0.26	0.42	0.678
Test [DSST] : Group2-1 * Time2-1	0.19	0.26	0.75	0.456
Test [SC] : Group2-1 * Time2-1	0.18	0.26	0.71	0.479
Test [SI] : Group2-1 * Time2-1	0.03	0.26	0.11	0.916
Random Effects				
σ2	0.57			
τ00Child	0.23			
ICC	0.29			
N Child	72			
Observations	705			
Marginal R2 / Conditional R2	0.072 / 0.344			

TMTa = trail making test version A, TMTb = trail making test version B, DSST = digit symbol substitution test, SC = Simon task congruent condition, SI = Simon task incongruent condition, Est. = estimate; SE = standar error

3.4 Intervention model 2: Pooled intervention period

3.4.1 Physical fitness

3.4.1.1 Model fitting

Model selection is documented in the Appendix Section 8.3.3.1. For the analysis of intervention effects of physical fitness, Group (pooled) and Time (pooled) nested under the factor levels of Test. In addition, BMI was included in the nested structure of Test.

3.4.1.2 Results

LMM revealed significant overall Time-related (pooled) improvements for 6 min run (β = .2; SE = .08; z = 2.54; p = .011) and star run (β = .27; SE = .08; z = 3.45; p = .001). For 6 min run (β = -.09; SE = .02; z = -5.59; p < .001), star run (β = -.05; SE = .02; z = -3.18; p = .002), and standing long jump (β = -.04; SE = .02; z = -2.5; p = .013), the LMM revealed lower BMI to be associated with better performances. However, significant interactions of Time (pooled) and Group (pooled) were not found for any of the implemented Tests. Results are presented in Table 3.3 and depicted in Figure 3.4. Means and standard deviations are reported in the Appendix Section 8.3.3.2.

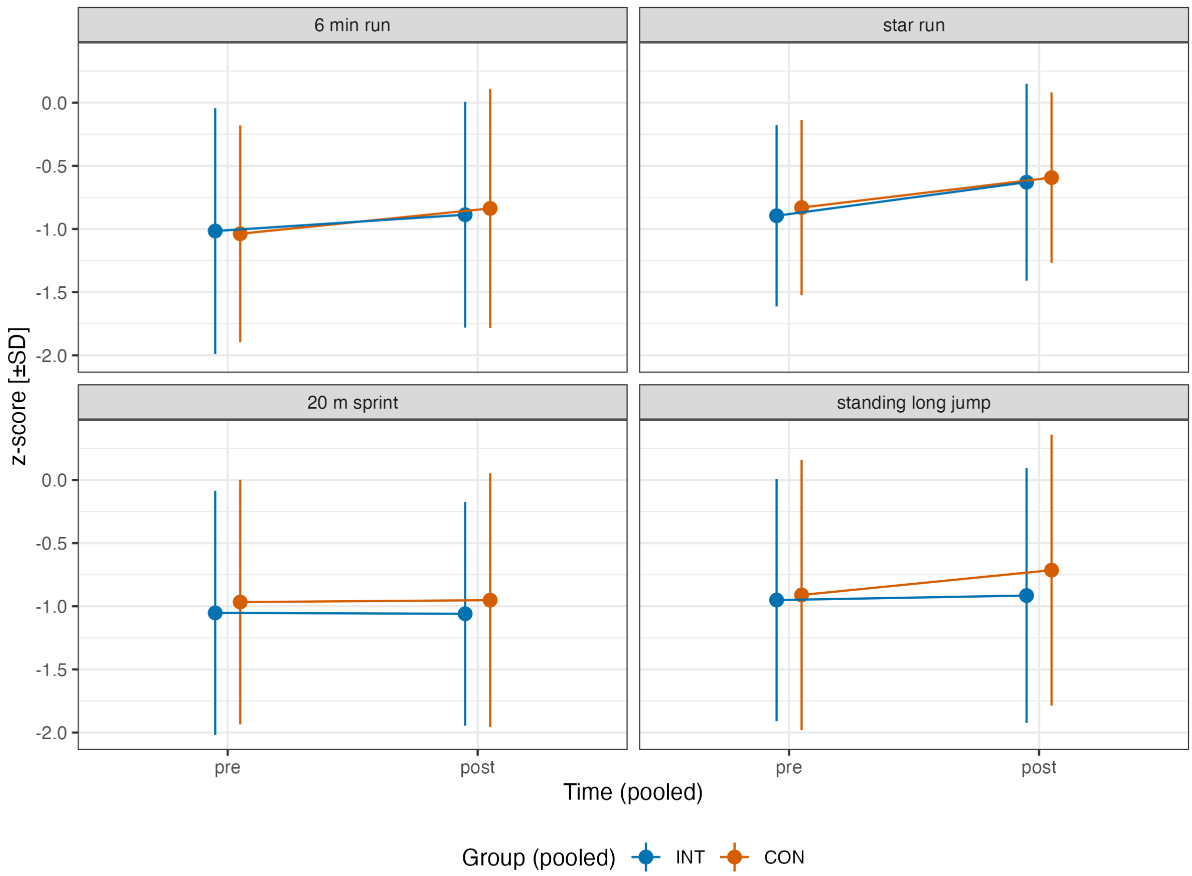


Figure 3. 4: Physical fitness plots for the pooled intervention period, CON = control group, INT = intervention group

Table 3. 3: Results of LMM for physical fitness for the pooled intervention period

Predictors	Est.	SE	z	P
Test [Run]	-0.94	0.08	-11.35	**<0.001**
Test [Star_r]	-0.73	0.08	-8.85	**<0.001**
Test [S20_r]	-1.00	0.08	-12.13	**<0.001**
Test [SLJ]	-0.87	0.08	-10.51	**<0.001**
Test [Run] : Group_pooled2-1	0.02	0.08	0.26	0.793
Test [Star_r] : Group_pooled2-1	-0.04	0.08	-0.46	0.645
Test [S20_r] : Group_pooled2-1	-0.07	0.08	-0.97	0.333
Test [SLJ] : Group_pooled2-1	-0.10	0.08	-1.33	0.184
Test [Run] : Time_pooled2-1	0.20	0.08	2.54	**0.011**
Test [Star_r] : Time_pooled2-1	0.27	0.08	3.45	**0.001**
Test [S20_r] : Time_pooled2-1	0.04	0.08	0.53	0.597
Test [SLJ] : Time_pooled2-1	0.15	0.08	1.92	0.055
Test [Run] : bmi_c	-0.09	0.02	-5.59	**<0.001**
Test [Star_r] : bmi_c	0.00	0.02	0.15	0.884
Test [S20_r] : bmi_c	-0.05	0.02	-3.18	**0.002**
Test [SLJ] : bmi_c	-0.04	0.02	-2.50	**0.013**
Test [Run] : Group_pooled2-1 * Time_pooled2-1	0.03	0.15	0.16	0.871
Test [Star_r] : Group_pooled2-1 * Time_pooled2-1	0.06	0.15	0.39	0.695
Test [S20_r] : Group_pooled2-1 * Time_pooled2-1	0.02	0.15	0.11	0.910
Test [SLJ] : Group_pooled2-1 * Time_pooled2-1	-0.13	0.15	-0.83	0.407
Random Effects				
σ2	0.37			
τ00Child	0.35			
ICC	0.49			
N Child	66			
Observations	1019			
Marginal R2 / Conditional R2	0.106 / 0.540			

Run = 6 min run, Star_r = star run, S20_r = 20 m sprint, SLJ = standing long jump, bmi_c = gender-specific zero-centered body mass index, Est. = estimate; SE = standar error

3.4.2 Executive function

3.4.2.1 Model fitting

Model selection is documented in the Appendix Section 8.3.4.1. For the analysis of intervention effects on executive function, Group (pooled) and Time (pooled) were nested within the factor levels of Test. In addition, age was also included in the fixed effects structure.

3.4.2.2 Results

LMM revealed overall age-related improvements in executive function (β = .89; SE = .07; z = 13.08; p < .001). However, no evidence for significant effects of Time (pooled), Group (pooled), or interactions of Time (pooled) and Group (pooled) were found for any of the implemented Tests. Results are presented in Table 3.4 and depicted in Figure 3.5. Means and standard deviations are reported in the Appendix Section 8.3.4.2.

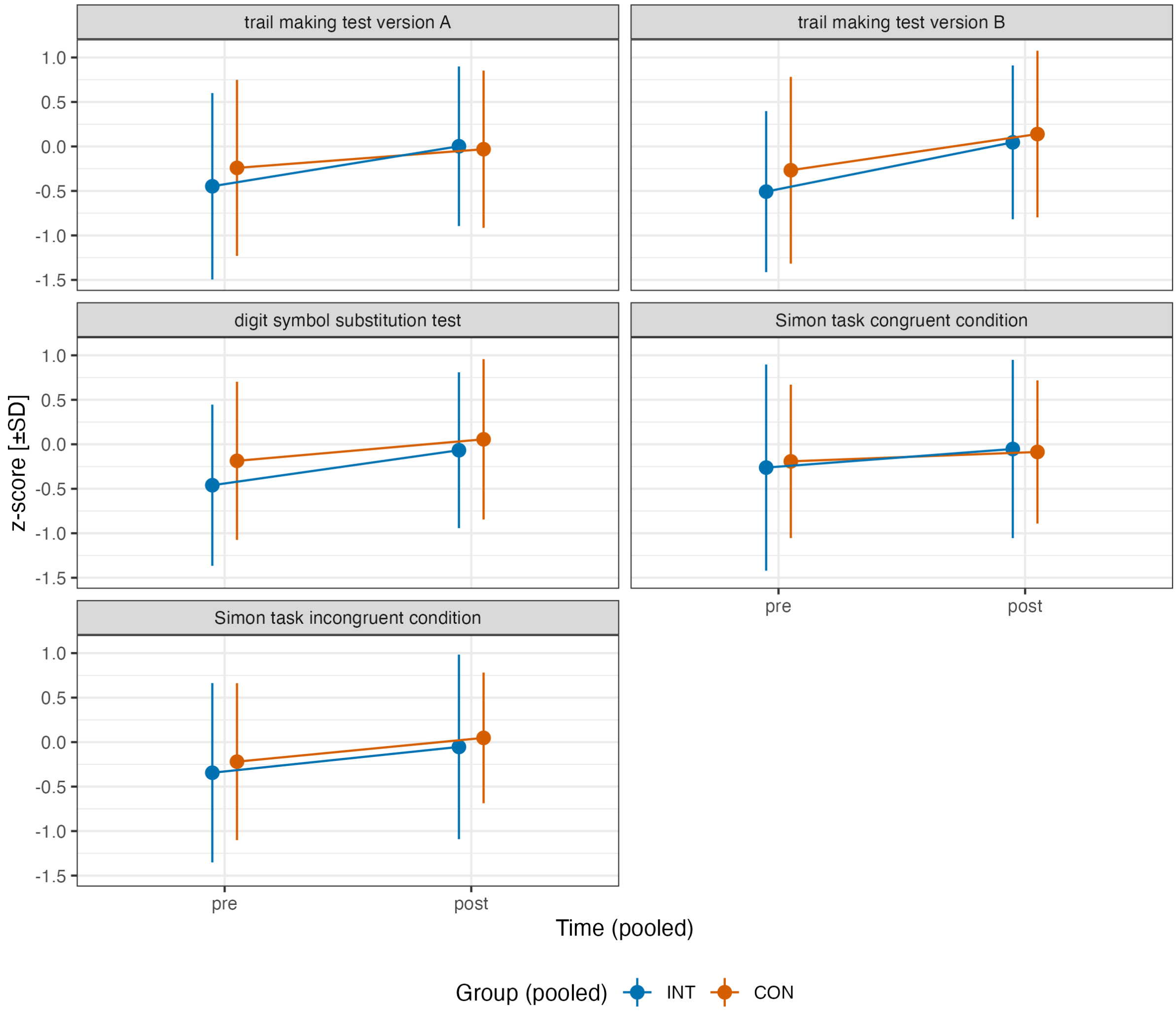


Figure 3. 5: Executive function plots for the pooled intervention period, CON = control group, INT = intervention group

Table 3. 4: Results of the LMM for executive function performance for the pooled intervention period

Predictors	Est.	SE	z	p
Test [TMTa]	-0.16	0.10	-1.68	0.093
Test [TMTb]	-0.13	0.10	-1.35	0.177
Test [DSST]	-0.15	0.10	-1.53	0.127
Test [SC]	-0.13	0.10	-1.35	0.178
Test [SI]	-0.12	0.10	-1.29	0.198
a1	0.89	0.07	13.08	**<0.001**
Test [TMTa] : Group_pooled2-1	-0.04	0.09	-0.45	0.649
Test [TMTb] : Group_pooled2-1	-0.12	0.09	-1.33	0.185
Test [DSST] : Group_pooled2-1	-0.15	0.09	-1.67	0.095
Test [SC] : Group_pooled2-1	0.02	0.09	0.25	0.804
Test [SI] : Group_pooled2-1	-0.07	0.09	-0.78	0.433
Test [TMTa] : Time_pooled2-1	0.00	0.09	0.00	0.996
Test [TMTb] : Time_pooled2-1	0.15	0.09	1.59	0.112
Test [DSST] : Time_pooled2-1	-0.01	0.09	-0.12	0.902
Test [SC] : Time_pooled2-1	-0.17	0.09	-1.79	0.074
Test [SI] : Time_pooled2-1	-0.05	0.09	-0.50	0.614
Test [TMTa] : Group_pooled2-1 * Time_pooled2-1	0.22	0.18	1.21	0.228
Test [TMTb] : Group_pooled2-1 * Time_pooled2-1	0.13	0.18	0.70	0.483
Test [DSST] : Group_pooled2-1 * Time_pooled2-1	0.13	0.18	0.73	0.468
Test [SC] : Group_pooled2-1 * Time_pooled2-1	0.08	0.18	0.42	0.671
Test [SI] : Group_pooled2-1 * Time_pooled2-1	-0.00	0.18	-0.02	0.988
Random Effects				
σ2	0.54			
τ00Child	0.48			
ICC	0.47			
N Child	66			
Observations	1289			
Marginal R2 / Conditional R2	0.238 / 0.594			

TMTa = trail making test version A, TMTb = trail making test version B, DSST = digit symbol substitution test, SC = Simon task congruent condition, SI = Simon task incongruent condition, a1 = zero-centered age, Est. = estimate; SE = standar error

3.5 Intervention model 3: Long-term effects

3.5.1 Physical fitness

3.5.1.1 Model fitting

Model selection is documented in the Appendix Section 8.3.5.1. For the analysis of long-term intervention effects on physical fitness, Group and Time were included in the fixed factor structure, which was nested within Test. In addition, BMI and Berlin proximity were included in the nested structure of Test.

3.5.1.2 Results

The LMM revealed a significant overall Time-related improvement for star run between t0 and t1 (β = .3; SE = .15; z = 2.07; p = .039) and t2 and t3 (β = .32; SE = .15; z = 2.17; p = .030), for 20 m sprint between t6 and t8 (β = .4; SE = .15; z = 2.58; p = .010), and for standing long jump between t6 and t8 (β = .5; SE = .15; z = 3.26; p = .001). Further, analysis revealed significant effects of BMI for 6 min run (β = -.09; SE = .02; z = -5.62; p < .001), and 20 m sprint (β = -.04; SE = .02; z = -2.78; p = .006) with better performances being associated with a lower BMI. Effects of Berlin proximity were found for 6 min run (β = .-81; SE = .3; z = -2.73; p = .006), and standing long jump (β = .-76; SE = .3; z = -2.56; p = .011) with better performances being associated with children living close to Berlin. However, no evidence for interactions of Time and Group were found for any of the implemented Tests. Results are presented in Table 3.5 and depicted in Figure 3.6. Means and standard deviations are reported in the Appendix Section 8.3.5.2.

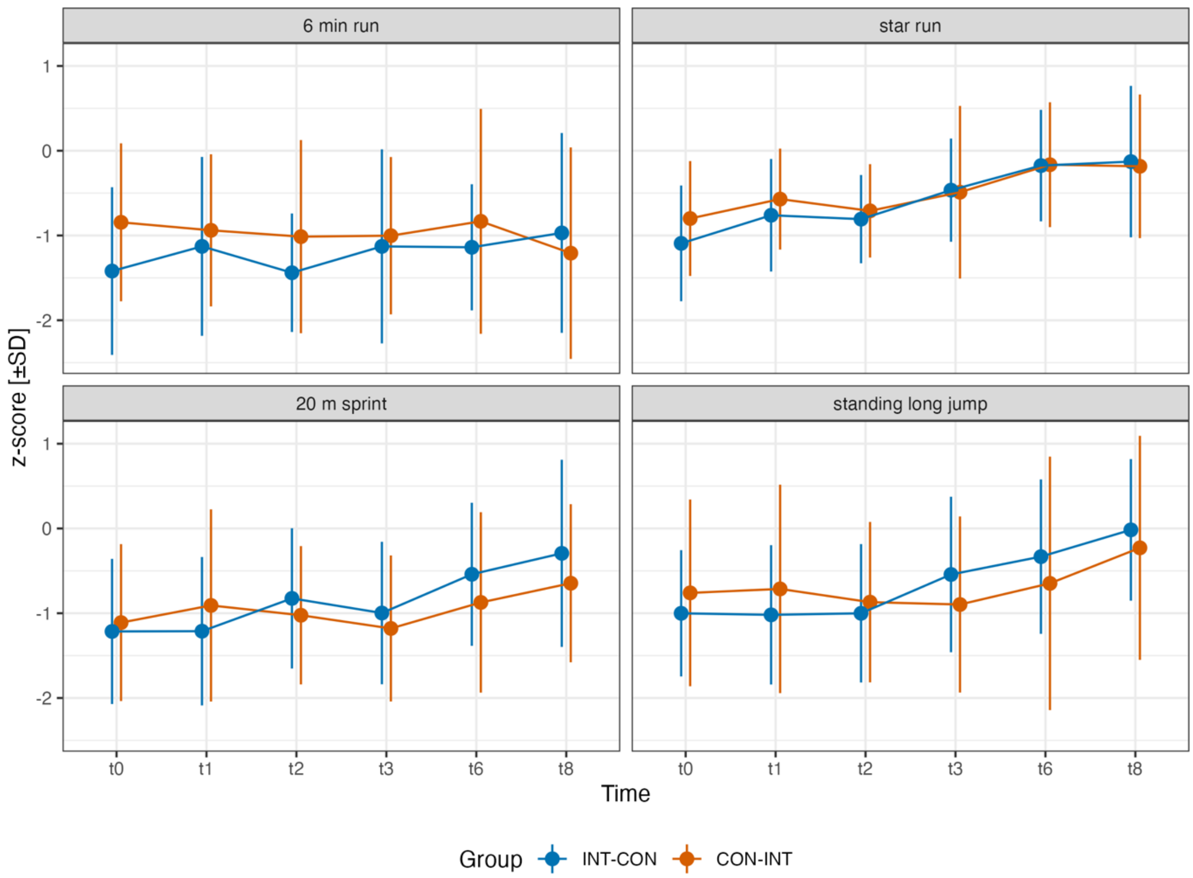


Figure 3. 6: : Physical fitness plots of long-term intervention effects, CON-INT = control condition between t0 and t1 and intervention condition between t2 and t3, INT-CON = intervention condition between t0 and t1 and control condition between t2 and t3

Table 3. 5: Results of LMM for physical fitness for the long-term intervention effects

Predictors	Est.	SE	z	P
Test [Run]	-0.88	0.13	-6.87	**<0.001**
Test [Star_r]	-0.45	0.13	-3.50	**<0.001**
Test [S20_r]	-0.74	0.13	-5.79	**<0.001**
Test [SLJ]	-0.46	0.13	-3.58	**<0.001**
Test [Run] : Group2-1	0.20	0.23	0.87	0.387
Test [Star_r] : Group2-1	0.01	0.23	0.03	0.974
Test [S20_r] : Group2-1	0.39	0.23	1.66	0.097
Test [SLJ] : Group2-1	0.37	0.23	1.61	0.108
Test [Run] : Time2-1	0.14	0.15	0.97	0.332
Test [Star_r] : Time2-1	0.30	0.15	2.07	**0.039**
Test [S20_r] : Time2-1	0.13	0.15	0.89	0.376
Test [SLJ] : Time2-1	0.03	0.15	0.17	0.862
Test [Run] : Time3-2	-0.16	0.15	-1.08	0.281
Test [Star_r] : Time3-2	-0.11	0.15	-0.75	0.452
Test [S20_r] : Time3-2	0.13	0.15	0.89	0.372
Test [SLJ] : Time3-2	-0.07	0.15	-0.45	0.652
Test [Run] : Time4-3	0.16	0.15	1.06	0.290
Test [Star_r] : Time4-3	0.32	0.15	2.17	**0.030**
Test [S20_r] : Time4-3	-0.11	0.15	-0.76	0.447
Test [SLJ] : Time4-3	0.26	0.15	1.79	0.074
Test [Run] : Time5-4	-0.02	0.15	-0.16	0.876
Test [Star_r] : Time5-4	0.26	0.15	1.72	0.087
Test [S20_r] : Time5-4	0.29	0.15	1.86	0.063
Test [SLJ] : Time5-4	0.14	0.15	0.89	0.374
Test [Run] : Time6-5	0.15	0.16	0.94	0.346
Test [Star_r] : Time6-5	0.07	0.15	0.45	0.651
Test [S20_r] : Time6-5	0.40	0.15	2.58	**0.010**
Test [SLJ] : Time6-5	0.50	0.15	3.26	**0.001**
Test [Run] : bmi_c	-0.09	0.02	-5.62	**<0.001**
Test [Star_r] : bmi_c	-0.01	0.02	-0.59	0.558
Test [S20_r] : bmi_c	-0.04	0.02	-2.78	**0.006**
Test [SLJ] : bmi_c	-0.03	0.02	-1.73	0.084
Test [Run] : bp2-1	-0.81	0.30	-2.73	**0.007**
Test [Star_r] : bp2-1	-0.25	0.30	-0.85	0.393
Test [S20_r] : bp2-1	-0.58	0.30	-1.95	0.052
Test [SLJ] : bp2-1	-0.76	0.30	-2.56	**0.011**
Test [Run] : Group2-1 * Time2-1	0.49	0.29	1.68	0.094
Test [Star_r] : Group2-1 * Time2-1	0.14	0.29	0.49	0.624
Test [S20_r] : Group2-1 * Time2-1	-0.13	0.29	-0.46	0.646
Test [SLJ] : Group2-1 * Time2-1	-0.02	0.29	-0.08	0.940
Test [Run] : Group2-1 * Time3-2	-0.31	0.30	-1.05	0.293
Test [Star_r] : Group2-1 * Time3-2	0.05	0.29	0.16	0.870
Test [S20_r] : Group2-1 * Time3-2	0.42	0.29	1.44	0.149

Table 3. 5: Results of LMM for physical fitness for the long-term intervention effects (continued

Predictors	Est.	SE	z	P
Test [SLJ] : Group2-1 * Time3-2	0.12	0.29	0.42	0.674
Test [Run] : Group2-1 * Time4-3	0.04	0.30	0.15	0.881
Test [Star_r] : Group2-1 * Time4-3	0.05	0.29	0.19	0.853
Test [S20_r] : Group2-1 * Time4-3	-0.10	0.29	-0.34	0.735
Test [SLJ] : Group2-1 * Time4-3	0.41	0.29	1.40	0.163
Test [Run] : Group2-1 * Time5-4	0.01	0.31	0.04	0.964
Test [Star_r] : Group2-1 * Time5-4	-0.07	0.31	-0.23	0.816
Test [S20_r] : Group2-1 * Time5-4	0.18	0.31	0.60	0.551
Test [SLJ] : Group2-1 * Time5-4	0.01	0.31	0.02	0.986
Test [Run] : Group2-1 * Time6-5	0.47	0.31	1.53	0.128
Test [Star_r] : Group2-1 * Time6-5	0.13	0.31	0.41	0.679
Test [S20_r] : Group2-1 * Time6-5	-0.00	0.30	-0.01	0.989
Test [SLJ] : Group2-1 * Time6-5	-0.15	0.30	-0.48	0.635
Random Effects				
σ2	0.38			
τ00Child	0.28			
ICC	0.43			
N Child	37			
Observations	836			
Marginal R2 / Conditional R2	0.313 / 0.605			

Run = 6 min run, Star_r = star run, S20_r = 20 m sprint, SLJ = standing long jump, bmi_c = gender-specific zero-centered body mass index, Est. = estimate; SE = standar error

3.5.2 Executive function

3.5.2.1 Model fitting

Model selection is documented in the Appendix Section 8.3.6.1. For the analysis of intervention on executive function, Group and Time were included in the fixed factor structure within Test. No other covariates were included in the LMM.

3.5.2.2 Results

The LMM revealed a significant overall Time-related improvement for trail making test version A between t0 and t1 (β = .55; SE = .16; z = 3.36; p = .001), t3 and t6 (β = .43; SE = .17; z = 2.53; p = .012), and t6 and t8 (β = .40; SE = .17; z = 2.37; p = .018). For trail making test version B, performance improved between t0 and t1 (β = .37; SE = .16; z = 2.27; p = .023), and t2 and t3 (β = .49; SE = .16; z = 3.01; p = .003). For digit symbol substitution test, performance improved between t1 and t2 (β = .37; SE=.16;z=2.25;p=.025),andt6andt8(β=.64;SE=.17;z=3.81;p<.001). Similarly,forSimon task congruent condition, performance improved between t1 and t2 (β = .44; SE = .16; z = 2.70; p

= .007), and t6 and t8 (β = .57; SE = .17; z = 3.37; p = .001); and for Simon task incongruent condition, performance improved between t0 and t1 (β = .34; SE = .16; z = 2.06; p = .039), and t6 and t8 (β = .75; SE = .17; z = 4.42; p < .001). However, there was no evidence for interactions of Time and Group for any of the implemented executive function Tests. Results are presented in Table 3.6 and depicted in Figure 3.7. Means and standard deviations are reported in the Appendix Section 8.3.6.2.

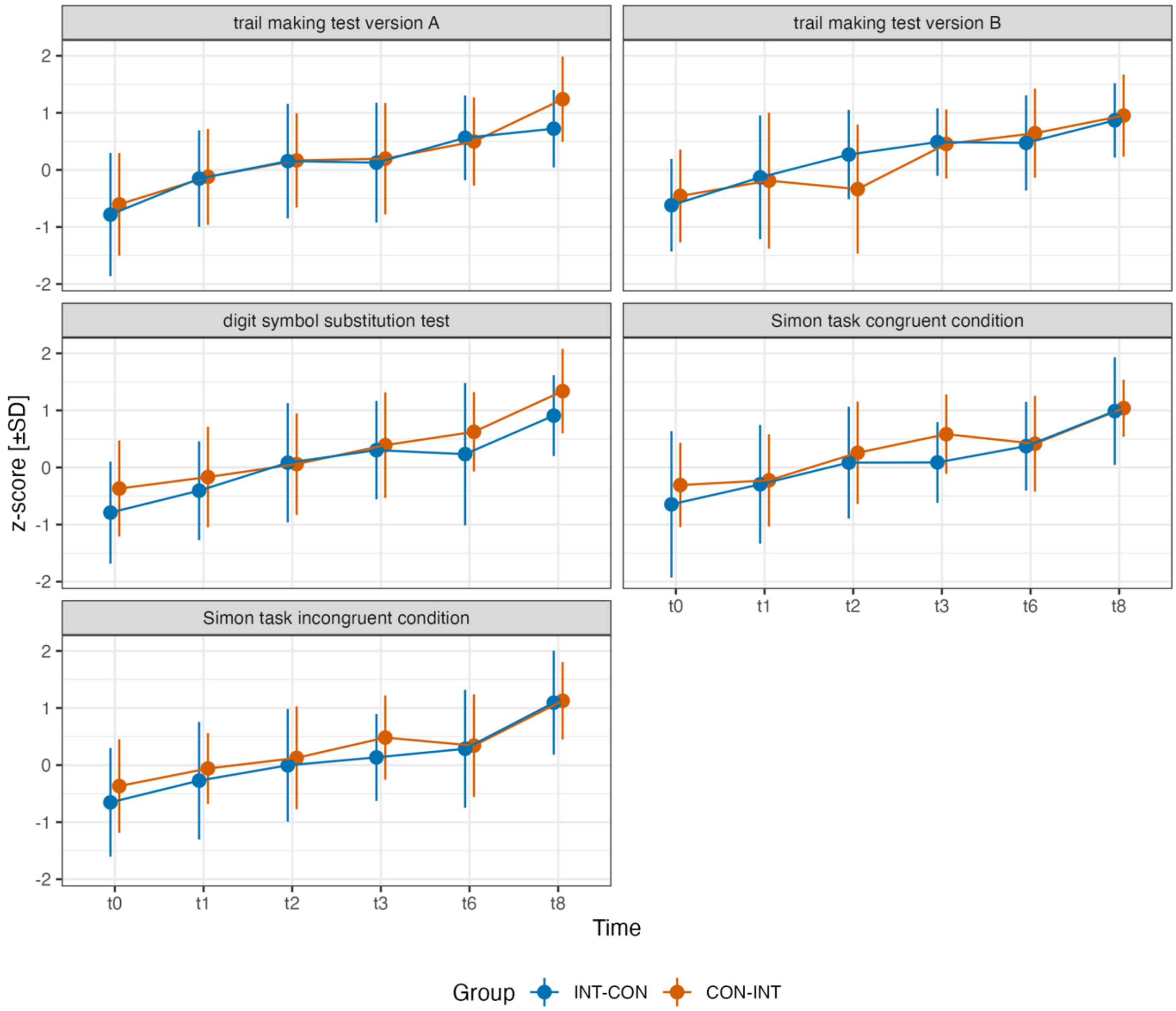


Figure 3. 7: Executive function performance plots of long-term intervention effects, CON-INT = control condition between t0 and t1 and intervention condition between t2 and t3, INT-CON = intervention condition between t0 and t1 and control condition between t2 and t3

Table 3. 6: Results of LMM for executive function performance for the long-term intervention effects

Predictors	Est.	SE	Z	p
Test [TMTa]	0.17	0.09	1.82	0.069
Test [TMTb]	0.21	0.09	2.18	**0.030**
Test [DSST]	0.19	0.09	2.00	**0.046**
Test [SC]	0.20	0.09	2.15	**0.032**
Test [SI]	0.19	0.09	2.03	**0.042**
Test [TMTa] : Group2-1	-0.14	0.19	-0.73	0.468
Test [TMTb] : Group2-1	0.03	0.19	0.18	0.860
Test [DSST] : Group2-1	-0.27	0.19	-1.43	0.153
Test [SC] : Group2-1	-0.21	0.19	-1.12	0.263
Test [SI] : Group2-1	-0.19	0.19	-1.03	0.305
Test [TMTa] : Time2-1	0.55	0.16	3.36	**0.001**
Test [TMTb] : Time2-1	0.37	0.16	2.27	**0.023**
Test [DSST] : Time2-1	0.28	0.16	1.74	0.081
Test [SC] : Time2-1	0.21	0.16	1.28	0.202
Test [SI] : Time2-1	0.34	0.16	2.06	**0.039**
Test [TMTa] : Time3-2	0.31	0.16	1.88	0.061
Test [TMTb] : Time3-2	0.13	0.16	0.82	0.411
Test [DSST] : Time3-2	0.37	0.16	2.25	**0.025**
Test [SC] : Time3-2	0.44	0.16	2.70	**0.007**
Test [SI] : Time3-2	0.24	0.16	1.45	0.148
Test [TMTa] : Time4-3	-0.01	0.16	-0.08	0.938
Test [TMTb] : Time4-3	0.49	0.16	3.01	**0.003**
Test [DSST] : Time4-3	0.26	0.16	1.60	0.109
Test [SC] : Time4-3	0.16	0.16	0.98	0.325
Test [SI] : Time4-3	0.24	0.16	1.50	0.135
Test [TMTa] : Time5-4	0.43	0.17	2.53	**0.012**
Test [TMTb] : Time5-4	0.15	0.17	0.89	0.373
Test [DSST] : Time5-4	0.15	0.17	0.86	0.389
Test [SC] : Time5-4	0.11	0.17	0.68	0.499
Test [SI] : Time5-4	0.06	0.17	0.36	0.719
Test [TMTa] : Time6-5	0.40	0.17	2.37	**0.018**
Test [TMTb] : Time6-5	0.30	0.17	1.79	0.075
Test [DSST] : Time6-5	0.64	0.17	3.81	**<0.001**
Test [SC] : Time6-5	0.57	0.17	3.37	**0.001**
Test [SI] : Time6-5	0.75	0.17	4.42	**<0.001**
Test [TMTa] : Group2-1 * Time2-1	0.08	0.33	0.25	0.806
Test [TMTb] : Group2-1 * Time2-1	0.15	0.33	0.47	0.637
Test [DSST] : Group2-1 * Time2-1	0.11	0.33	0.35	0.727
Test [SC] : Group2-1 * Time2-1	0.20	0.33	0.61	0.539
Test [SI] : Group2-1 * Time2-1	0.01	0.33	0.02	0.987
Test [TMTa] : Group2-1 * Time3-2	0.09	0.33	0.27	0.789
Test [TMTb] : Group2-1 * Time3-2	0.61	0.33	1.89	0.060
Test [DSST] : Group2-1 * Time3-2	0.33	0.33	1.02	0.309

Table 3. 6: Results of LMM for executive function performance for the long-term intervention effects (continued)

Predictors	Est.	SE	Z	p
Test [SC] : Group2-1 * Time3-2	-0.03	0.33	-0.11	0.915
Test [SI] : Group2-1 * Time3-2	0.15	0.33	0.46	0.644
Test [TMTa] : Group2-1 * Time4-3	-0.03	0.33	-0.09	0.931
Test [TMTb] : Group2-1 * Time4-3	-0.54	0.33	-1.66	0.097
Test [DSST] : Group2-1 * Time4-3	-0.08	0.33	-0.25	0.802
Test [SC] : Group2-1 * Time4-3	-0.31	0.32	-0.96	0.335
Test [SI] : Group2-1 * Time4-3	-0.21	0.32	-0.64	0.524
Test [TMTa] : Group2-1 * Time5-4	0.06	0.34	0.19	0.852
Test [TMTb] : Group2-1 * Time5-4	-0.27	0.34	-0.80	0.423
Test [DSST] : Group2-1 * Time5-4	-0.38	0.34	-1.10	0.273
Test [SC] : Group2-1 * Time5-4	0.40	0.34	1.17	0.241
Test [SI] : Group2-1 * Time5-4	0.24	0.34	0.70	0.483
Test [TMTa] : Group2-1 * Time6-5	-0.54	0.34	-1.60	0.110
Test [TMTb] : Group2-1 * Time6-5	0.13	0.34	0.38	0.706
Test [DSST] : Group2-1 * Time6-5	0.00	0.34	0.01	0.991
Test [SC] : Group2-1 * Time6-5	0.03	0.34	0.10	0.921
Test [SI] : Group2-1 * Time6-5	0.06	0.34	0.19	0.849
Random Effects				
σ2	0.48			
τ00Child	0.25			
ICC	0.34			
N Child	37			
Observations	1062			
Marginal R2 / Conditional R2	0.285 / 0.529			

TMTa = trail making test version a, TMTb = trail making test version b, DSST = digit symbol substitution test, SC = Simon task congruent condition, SI = Simon task incongruent condition, Est. = estimate; SE = standar error

3.6 Summary of results

The most striking outcome of assessing the effects of the implemented 14-week quantitative remedial physical education intervention on physical fitness and executive function of children with deficits in their physical fitness, was (1) the absence of detectable intervention-related improvements of physical fitness or executive functions in any of the approaches used.

Other results revealed by the LMMs were (2) several improvements over time in all tests of physical fitness and executive functions, (3) interactions between BMI and 6 min run, star run, and standing long jump performances, with better performances in children with a lower BMI, (4) an effect of proximity to Berlin on performance in 6 min run and standing long jump, with children living closer to Berlin performing better, (5) improvements in executive function test performance with age, and (6) gendered differences in executive function test performances, with boys performing better compared to girls.

3.7 Discussion

Regarding the absence of detectable intervention effects, two potential factors hold plausible explanations in relation to the remedial physical education intervention: (1) duration of the intervention period and (2) exercise specificity. A meta-analysis summarising effects of additional physical education found positive effects in cardiorespiratory and muscular fitness, however the study included predominantly studies with intervention periods lasting one year or longer (i.e., 17 out of 20 studies) (García-Hermoso et al., 2020). One of the included studies examined effects of two additional physical education lessons over a 16-week period in adolescents aged 12 to 14 years and found intervention-related effects for shuttle run but not standing long jump and 4 x 10 m shuttle run performances (Ardoy et al., 2011). In another study, additional physical education lessons were implemented for 8 weeks, however, the publication was only available in Spanish and therefore could not be further discussed in this thesis (Lechuga et al., 2012). This could indicate that intervention periods of 14 weeks, as used in the SMaRTER study, might be insufficient to achieve detectable intervention effects.

Studies examining effects of more specific exercises (i.e., exercises that focus on improving specific subcomponents of physical fitness) integrated into existing physical education classes have shown to induce intervention-related benefits compared to regular physical education lessons. For example, studies in which 9- to 10-week plyometric programs were integrated into the physical education curriculum with children aged 8 to 12 years found greater improvements in sprint speed, vertical jump height, standing long jump distance, number of push-ups, and half-mile run performance compared to regular physical education controls[97] (Faigenbaum & Farrell, 2009; Kotzamanidis, 2006). Similarly, in a study examining effects of an 8-week resistance training program incorporated into physical education in children aged 10 to 12 years, greater intervention-related improvements were found in sit-up and arm bend performances, but not in standing long jump distance, compared to regular physical education[98] (Viciana et al., 2013). As the SMaRTER study included a comprehensive intervention program aimed at improving several components of physical fitness in children, the respective exercises might have been insufficient to induce intervention-related improvements in physical fitness components. While a 14-week intervention curriculum aimed at improving a specific component of physical fitness could have achieved significant intervention effects, the SMaRTER study was oriented towards bringing general physical fitness of participating children to a level comparable to that of their peers, while offering exercises in a curriculum that was fun and provided a safe community learning experience for participating children (KMK, 1999). Accordingly, implementing the SMaRTER physical education curriculum over a longer period of time (i.e., ≥ one school year) may have resulted in measurable intervention-related improvements in physical fitness and executive function.

A similar dynamic can be observed in studies examining effects of physical education-based interventions on executive function. For example, a review article examining effects of additional physical education on executive function and academic performance predominantly included studies with intervention durations longer than one year (i.e., 3 out of 4 studies) (Garciá-Hermoso et al., 2021). Shorter interventions based on physical education tend to implement training protocols involving some form of high-intensity exercise such as high-intensity interval training (Costigan et al., 2016; Garciá-Hermoso et al., 2021; Martínez-Lopez et al., 2018; Xue et al., 2019). Thus, similarly to the lack of detectable intervention-related improvements on physical fitness, the lack of intervention-related effects on executive function in the SMaRTER study may also have attributed to an insufficient intervention period duration or an insufficient training intensity.

Another factor to consider in this discussion would be the inclusion of children with deficits in their physical fitness. In the SMaRTER study, it was hypothesised that children with deficits in their physical fitness would make greater improvements when exposed to additional exercise due to their larger potentials. However, this view is biased by the assumption that deficits in physical fitness are caused by a mere lack of exercise or can at least be properly addressed by additional exercising, thereby neglecting, for example, sociocultural conditions related to children's physical fitness levels. This will be discussed further in Section 7.3.

However, in addition to the absence of detectable intervention-related effects, the different analyses found several time-related improvements of physical fitness and executive function, as well as moderating effects of BMI, gender, and proximity to Berlin. The various significant results of temporal improvements in physical fitness and executive function tests, as well as age-related improvements in executive function, are consistent with expectations arising from research on growth- and maturation-related changes in children (see Section 1.1.2 for more details on age-/growth-/maturation-related changes in physical fitness; details on age/growth/maturity-related changes in cognition can be found elsewhere (Diamond, 2000, 2006; Jacobsen et al., 2017; S. C. Li, 2003; Thompson & Steinbeis, 2020)).

In preliminary analysis of baseline differences (described in Section 2.7), interactions were found between group and gender and several anthropometric variables. Therefore, BMI was included in intervention analyses and showed moderation of performance in physical fitness but not executive function. These interactions further informed the analysis of body height and body mass effects on physical fitness, which are assessed in Chapter 4. Including BMI in the analyses of potential intervention-related effects showed that for 6 min run, 20 m sprint, and standing long jump, but not star run, better performances were associated with a lower BMI. These results are mostly consistent with the assumption that body weight-related parameters particularly influence performance in physical fitness tests that take total body weight into account Deforche et al. (2003). Star run performances have shown no association with BMI, although star run has strong construct correlations with the other three tests included in the analyses (Fühner et al., 2021) and also uses whole-body movements. This could be an indication that other elements such as cognitive elements of this task might be a stronger limiting factor and thus be less affiliated with anthropological parameters. This interpretation would further be supported as BMI did not reveal to be of any statistical relevance in the executive function models. However, as there is evidence of a relationship between BMI and coordinative motor tasks in children aged 6 to 11 years (Giuriato et al., 2019; Lopes et al., 2018), the relationship between BMI and star run performances might be different in children with deficits in their physical fitness compared to regular fit children.

Effects of gender on cognitive performance favoured boys in one of the three models, which contradicts some findings from studies that examined effects of gender on executive functions between the ages of 5 and 12 and found that they were either absent or in favour of girls (Ishihara et al., 2018; Jacobsen et al., 2017; Yamamoto & Imai-Matsumura, 2019). While it is possible that gender-specific distributions of executive functions are different for children with deficits in physical fitness compared to regularly fit children, evidence for such distributions is missing. However, as effects of gender on executive function only occurred in one of the three models, it might be a sample-specific or spurious result.

The model for evaluating mid- and long-term intervention effects showed a positive effect of proximity to Berlin on performances in 6 min run and standing long jump, with children living closer to Berlin performing better. These results are in line with unpublished results of the

EMOTIKON study, which show that children who live closer to Berlin performed better in 6 min run, 20 m sprint, and standing long jump (R. Kliegl & Teich, 2022). Since this model is the only one that includes data points after the onset of the covid pandemic, it is possible that the decline in physical fitness associated with the covid pandemic was different for children living in denser and less isolated environments than for children living in more rural areas. This assumption is in line with results of a study examining the impact of the covid pandemic on physical fitness of Slovenian 6th and 8th grade children, which found a greater decline in overall physical fitness, but not in 600 m running performance among rural youth compared to urban youth (Pajek, 2022). However, due to the small dropout-related sample size of the mid- and long-term effects model (n = 37; CON-INT = 18, INT-CON = 19), results could be spurious and should be interpreted with care.

3.8 Implications and subsequent analyses

The lack of detectable intervention-related effects on performance in the physical fitness tests used allows the distinction between intervention and control group children (i.e., the factor "group") to be omitted in the following chapters. Accordingly, these chapters will focus on effects of various moderators of physical fitness in children with motor deficits, as already outlined in Section 2.8. In addition, a prominent finding from the analysis of intervention effects on physical fitness components was, that including BMI in the analyses improved the models fits and revealed significant effects on 6 min run, 20 m sprint, and standing long jump with better performances being associated with lower body mass indices. This poses the question how the subcomponents of BMI (i.e., body height and body mass) are moderating physical fitness performances in children with deficits in their physical fitness. This will be assessed in the following chapters where the conceptualisation of physical fitness is expanded by including the ball push test.

4 Analyses of body composition-related effects on physical fitness

4.1 Introduction

The results of the Chapter 3 study show that body mass index (BMI) has a detectable negative effect on physical performances in 6 min run, 20 m sprint, and standing long jump, but not in the star run. However, using BMI as a composite of body height and body mass could potentially mask opposing effects of body height and body mass. Furthermore, including the interaction of both body parameters into the model would allow for a more dynamic understanding of how these parameters interact with the different components of physical fitness. Thus, the main focus of this chapter is on the effects of body height and body mass on physical fitness performance of children with deficits in their physical fitness.

4.2 Statistical analysis and model fitting

For the analysis of body mass and body height effects on physical fitness, 6 min run, 20 m sprint, star run, standing long jump, and ball push test were included in the analysis. As outlined in Section 2.8, ball push test was included in the conceptualisation of physical fitness. This decision was based on the hypothesis that a higher body mass might be beneficial for test performances as body weight could be used as a counterweight. As this hypothesis could not be applied to the one leg balance test, this test was not included in the hypothesis. For details on preprocessing and z-transformation of physical fitness data, see Section 8.1.4, and for anthropological parameters, see Section 8.1.3 of the Appendix. Missing data for body mass and body height were imputed from previous assessments.

A linear mixed models (LMM) analysis approach was implemented, and a parsimonious model selection approach was adopted (D. M. Bates et al., 2018). Z-scores of the above mentioned physical fitness tests were used and the physical fitness tests were set as factor levels of Test and successive difference contrasts were applied. Body mass and body height were included as fixed effects nested within the factor levels of Test. Additionally, age was included as a fixed effect. Child was included as a random factor with age as a variance component. Residual-based diagnostics (e.g., q-q plot, standardized residuals over fitted values, etc.) were carried out for the final model and reported in the results.

The process of model selection is documented in the Appendix Section 8.4.2. The final LMM was fitted with age and Test as fixed factors. Body height and linear and quadratic mass with interactions between height and linear mass were nested within Test. Child was entered as a random factor, including individual age as a variance component.

4.3 Results

4.3.1 Fixed effects

LMM showed a significant overall effect for age ($\beta = 14$; SE = .06; z = 2.5; $p = .013$) with performance increasing with age.

Regarding anthropometric parameters, the LMM revealed better performances for taller children in all included physical fitness tests (i.e., 6 min run ($\beta = .05$; SE = .01; z = 4.44; $p < .001$), star run ($\beta = .04$; SE = .01; z = 3.2; $p < .001$), 20 m sprint ($\beta = .04$; SE = .01; z = 3.78; $p = .001$), standing long jump ($\beta = .06$; SE = .01; z = 4.96; $p < .001$), and ball push test ($\beta = .05$; SE = .01; z = 4.53; $p < .001$).

For body mass, LMM revealed significant negative linear effects for 6 min run (β = -.05; SE = .01; z = -6.61; p < 0.001), and standing long jump (β = -.04; SE = .01; z = -5.48; p < 0.001), with increases in body mass being associated with a decrease in performance. A significant positive linear effect was found for ball push test (β = .03; SE = .01; z = 3.86; p < 0.001), with increases in body mass being associated with increases in ball push distance. The effect of body mass on 20 m sprint performance was characterized by a negative quadratic effect 20 m sprint (β = -.0; SE = .0; z = -2.52; p = .012), indicating that an increase of body mass was first associated with an increase of sprint performance, but then followed by a performance decrease.

For 6 min run (β = .00; SE = .00; z = 2.52; p = .012), and 20 m sprint (β = .00; SE = .00; z =2.30; p = .022), the LMM revealed a significant interaction of body height and body mass. Whereas in smaller children a higher body mass was associated with a lower performance, in taller children a higher body mass was associated with a better performance.

Model results are reported in Table 4.1. The depiction of significant effects is done for raw values as well as partial effects in Figure 4.1. Means and standard deviations of included variables are reported in the Appendix Section 8.4.3.

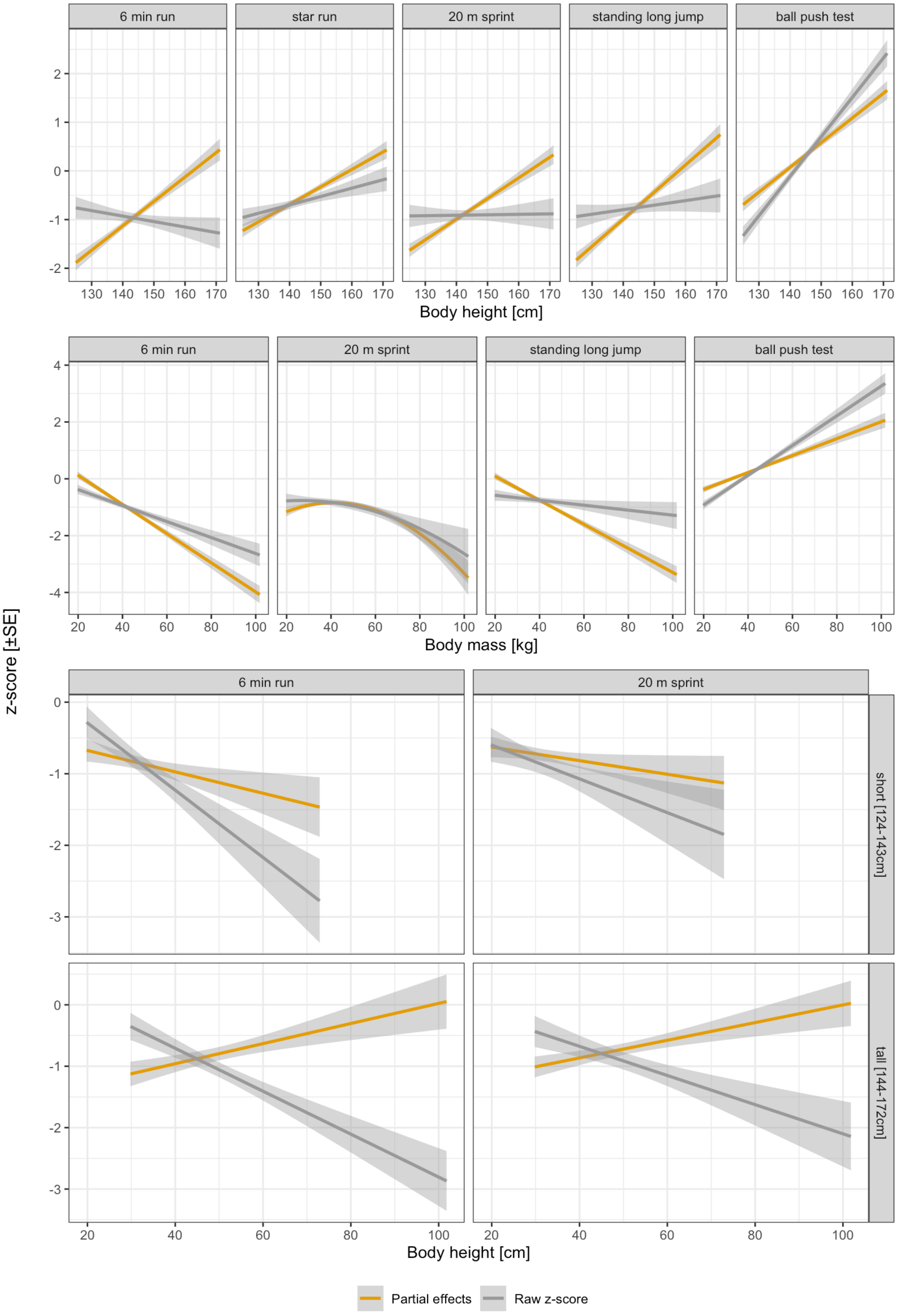


Figure 4. 1: Effects of body height, body mass, and the interaction of body height and body mass on physical fitness z-scores and partial effects; SE = standard error

Table 4. 1 : Effects of body height and body mass on physical performance

Predictors	Est.	SE	z	p
Test [Run]	-0.95	0.08	-11.54	**<0.001**
Test [Star_r]	-0.56	0.08	-6.81	**<0.001**
Test [S20_r]	-0.84	0.08	-10.22	**<0.001**
Test [SLJ]	-0.80	0.08	-9.66	**<0.001**
Test [BPT]	0.25	0.08	3.05	**0.002**
a1	0.14	0.06	2.50	**0.013**
Test [Run] : h	0.05	0.01	4.44	**<0.001**
Test [Star_r] : h	0.04	0.01	3.20	**0.001**
Test [S20_r] : h	0.04	0.01	3.78	**<0.001**
Test [SLJ] : h	0.06	0.01	4.96	**<0.001**
Test [BPT] : h	0.05	0.01	4.53	**<0.001**
Test [Run] : m	-0.05	0.01	-6.61	**<0.001**
Test [Star_r] : m	-0.01	0.01	-1.58	0.114
Test [S20_r] : m	-0.03	0.01	-4.12	**<0.001**
Test [SLJ] : m	-0.04	0.01	-5.48	**<0.001**
Test [BPT] : m	0.03	0.01	3.86	**<0.001**
Test [Run] : m2	-0.00	0.00	-1.91	0.056
Test [Star_r] : m2	-0.00	0.00	-1.57	0.117
Test [S20_r] : m2	-0.00	0.00	-2.52	**0.012**
Test [SLJ] : m2	0.00	0.00	1.36	0.174
Test [BPT] : m2	-0.00	0.00	-0.76	0.446
Test [Run] : h * m	0.00	0.00	2.52	**0.012**
Test [Star_r] : h * m	0.00	0.00	0.92	0.356
Test [S20_r] : h * m	0.00	0.00	2.30	**0.022**
Test [SLJ] : h * m	-0.00	0.00	-0.41	0.680
Test [BPT] : h * m	-0.00	0.00	-0.23	0.819
Random Effects				
σ2	0.38			
τ00Child	0.35			
τ11Child.a1	0.02			
ρ01Child	0.57			
ICC	0.49			
N Child	76			
Observations	1771			
Marginal R2 / Conditional R2	0.396 / 0.690			

Run = 6 min run, Star_r = star run, S20_r = 20 m sprint, SLJ = standing long jump, BPT = ball push test, a1 = zero centered age, h = linear zero centered body height, m = linear zero centered body mass, m2 = quadratic zero centered body massEst. = estimate; SE = standar error

4.3.2 Variance components and correlation parameters

The random effect structure revealed differences between children in age effects on physical fitness. The child-related grand mean positively correlated with age effects (r = .57), meaning that fitter children were more likely to show greater age-related improvements. Individual variance components and correlation parameters are reported in Table 4.2, and the correlation between grand mean and age-related effects is plotted in Figure 4.2.

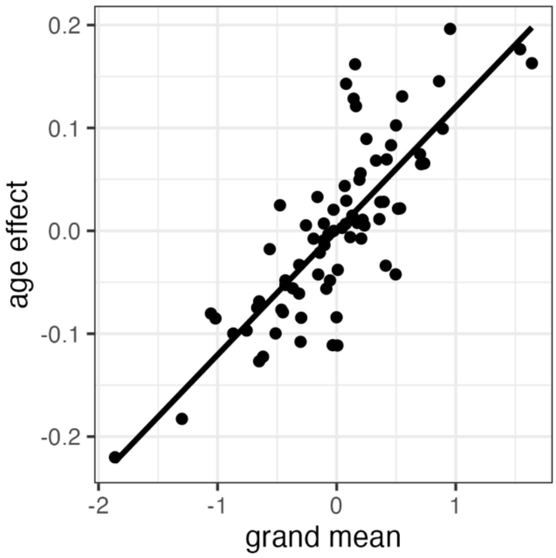


Figure 4. 2: Correlation of children's grand mean and age effects

Table 4. 2: Variance components and correlation parameter of 'm3_hm1'

Groups	Name	Variance	SD	Corr
Child	grand mean	0.345	0.588	
Child	a1	0.016	0.127	0.573
Residual		0.378	0.615	
a1 = age centered at 0				

4.3.3 Model diagnostics

Model performance parameters did not reveal any violations of LMM requirements and are plotted in Figure 4.3.

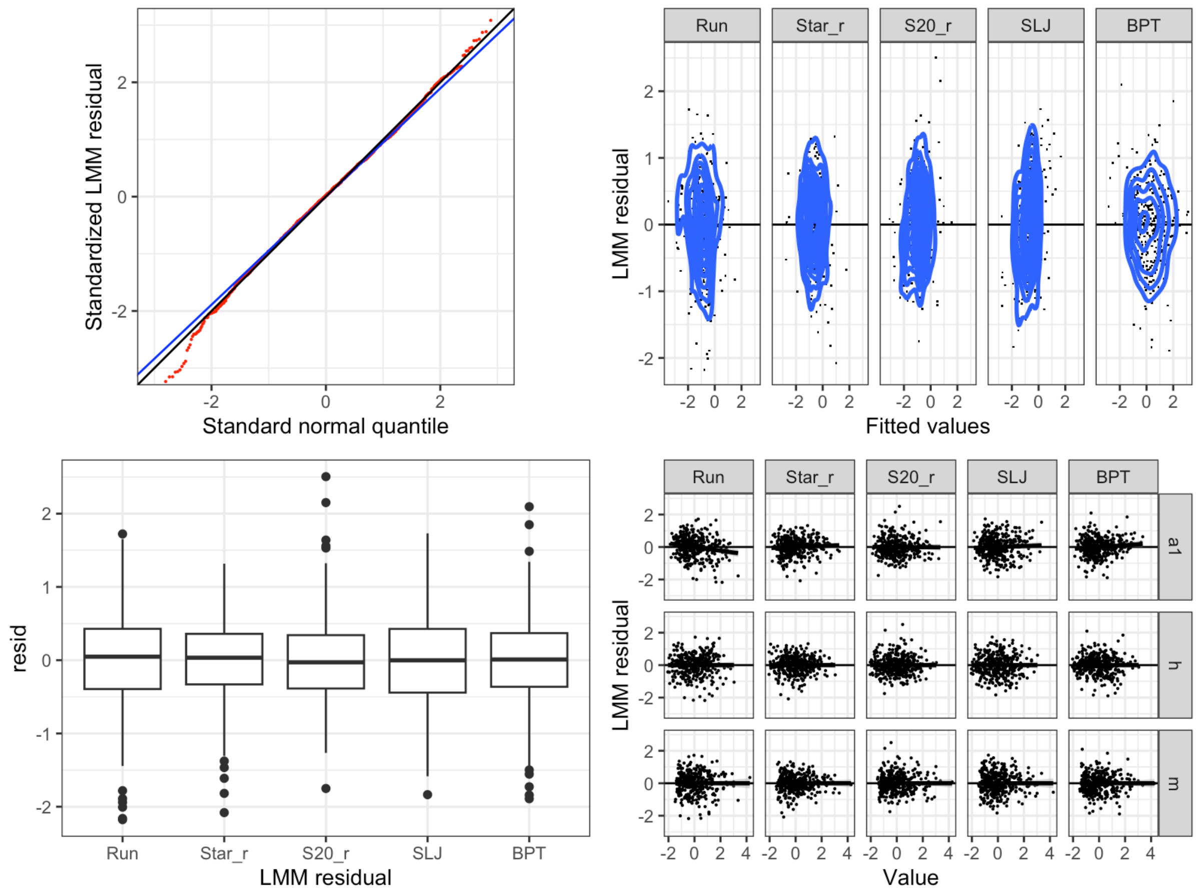


Figure 4. 3 : Model performance parameters

4.4 Summary of results

This LMM revealed that (1) physical performance increased with body height in all tests used, (2) negative linear effects of body mass were found for performances in 6 min run and standing long jump, with children with higher body mass showing lower test performances, (3) a negative quadratic trend was found for 20 m sprint performances, with stronger negative effects occurring with increasing body mass, (4) a positive effect was found of body mass on ball push test performances, with improvements in performances in children with higher levels of body mass, and (5) significant interactions between body height and body mass on 6 min run and 20 m sprint performances, with higher levels of body mass being associated with performance improvements in taller children, while in shorter children increases in body mass were associated with decreases in performance. Additionally, the LMM revealed (6) general age-related improvements in all tests of physical fitness, and (7) a positive correlation of individual variance components between overall physical fitness and age, with fitter children more likely to show greater age-related improvements.

4.5 Discussion

Similar trends for body height and body mass were found in studies that used allometric modelling to assess effects of body height and body mass on 20 m shuttle run, 12 min run, cooper test, 40 m sprint, 10 x 5 m shuttle run, vertical jump, standing long jump performances, and handgrip strength. The approach yielded predominantly positive associations for body height and negative associations for body mass with test performances, with the only exception being handgrip strength, where a positive association was found between body mass and test performances[99] (Giuriato et al., 2020; Lovecchio et al., 2019; Nevill et al., 2009; Silva et al., 2016; Valdivia et al., 2015). As argued in these studies, positive effects found in this chapter of body height on standing long jump, 6 min run, 20 m sprint, and star run performances can be attributed to physical advantages associated with an elevated center of gravity, longer legs, and slender stature (Giuriato et al., 2020; Lovecchio et al., 2019; M. A. M. dos Santos et al., 2018; Silva et al., 2016; Valdivia et al., 2015). Negative effects of body mass on standing long jump, 6 min run, and 20 m sprint performances can be attributed to excess body mass associated with body fat (Norman et al., 2005). In fact, introducing body fat percent into allometric analyses of the association of body height and mass with standing long jump, a negative association with standing long jump performance was found for body fat percent and a positive association for body mass (Lovecchio et al., 2019). Of note, indications of a negative effect were also found for body mass on star run performances ($\beta = -.01$; SE = .01; z = -1.58; p = .114; see Table 4.1). This trend was not significant in the present sample, but may reach significance in higher-powered studies with larger samples, especially as body mass has been shown to be negatively associated with gross motor coordination performances[100] (M. A. M. dos Santos et al., 2018). For ball push test performance, effects of body height and mass are similar to those reported for handgrip strength (Nevill et al., 2009; Silva et al., 2016; Valdivia et al., 2015). Results are consistent with the hypothesis that fitness tests that do not utilise total body mass benefit from increased muscle mass without being affected by fat mass (Deforche et al., 2003; Ervin et al., 2014). Furthermore, it has been theorised that in addition to more efficient use of muscle mass, body mass can act as leverage or counterweight to further increase strength performance, resulting in greater than expected improvements in the respective muscular strength tests (Asmussen & Heebøll-Nielsen, 1955; Nevill et al., 2009).

In addition, the LMM showed significant interactions between body height and body mass affecting performances in 6 min run and 20 m sprint. Here, taller children exhibit a positive relationship between body mass and test performances, while smaller children's performances decreased with increasing body mass. When interpreting these results, it should be taken into account that body mass in children increases with height by a power of approximately two during growth (Chung, 2015). Another study found that fat-free mass in children scales with height by a power of about three, while fat mass does not appear to scale with height (Z. Wang et al., 2012). This would suggest that body mass gains in larger children might be more skewed towards fat-free mass, while in smaller children the same absolute body mass gains would be more skewed towards fat mass. As there seems to be a positive correlation between fat-free mass and cardiorespiratory and muscular fitness, and a negative correlation for fat mass (Joensuu et al., 2018), this could explain the interaction found between body height and mass. That is, in larger children, mass gain is more likely to be associated with an increase in fat-free mass, which increases performances in 6 min run and 20 m sprint, while in smaller children, the same absolute mass gains are more likely to be associated with fat mass, which adversely affects performances in these tests. However, given scalings of body mass parameters with height cited above, it would be interesting to examine how interactions between different body size parameters relate to physical fitness test performances. More specifically, deconstructing the widely used BMI into its subcomponents and their interactions and assessing the relevance of these subcomponents and interactions for different components of physical fitness.

Aside anthropometric variables, age was included as a fixed effect and a variance component in the individual random effects. Including age in the fixed effects revealed overall expected age-related improvements in all test performances (Albrecht, 2015; Niessner et al., 2021; Ortega et al., 2023). Interestingly, including age as a variance component in the individual random factors revealed a significant correlation between overall physical fitness (i.e., grand mean) and age-related improvements, with fitter children being more likely to show greater age-related increases in physical fitness. Comparable results were found in analyses of the EMOTIKON study, where age of children was included as a variance component in the random effect structure of schools. Analysis of physical fitness of 108,295 third graders aged 8 to 8.99 years showed a positive correlation (r = .48) between grand mean and age-related effects at school level (Fühner et al., 2021). These results were replicated in the covid pandemic effects analysis in 83,476 third graders aged 8.05 to 9.11 years (r = .49) in the EMOTIKON sample (Teich et al., 2023). The authors argue that either schools that offer more fitness-promoting resources could produce greater gains in children's physical fitness, or fitter children could have greater fitness gains in a year, leading to greater overall school-related fitness as well as greater age-related fitness improvements (Fühner et al., 2021). The present study shows that the correlation is replicable at the individual level and in a smaller sample, and might therefore also exist within schools. It remains unclear whether this is driven by children using resources at their schools to promote fitness, or by children with greater age-related improvements showing a steeper trajectory of physical fitness over their lifetime, leading to higher levels of physical fitness when assessed within a specific time period. However, the correlations could also be indicative of socioeconomic dependencies of schools and children's families. Quality of schools is related to its areas socioeconomic status (Nieuwenhuis & Xu, 2021), meaning that "better" schools tend to be located in more affluent neighbourhoods, and that differences in within-school performance are related to the socioeconomic background of children's families (Ma, 2000). This might suggest that the correlation between overall physical fitness and age-related improvement in physical fitness, both at school and individual level, might be at least partly due to different opportunities available to each child depending on the socioeconomic background of the area and their family (Rittsteiger et al., 2021).

4.6 Implications and subsequent analyses

Considering the overall positive effects of body height as well as age on physical fitness performances, the question remains whether these effects are associated with growth and maturation of the included children. Further, regarding the positive correlation between the overall level of physical fitness and the magnitude of age effects on an individual level, it remains unclear how these effects are moderated. Accordingly, the following Chapter 5 will assess the effects of age and the maturity offset estimated using two different formulas (Mirwald et al., 2002; Moore et al., 2015) on physical fitness. Of note, as the ball push test might be differently associated with changes in physical fitness associated with growth and age, it was not included in the following analysis to allow a more precise modeling of the effects of age and maturity offset and a latent construct of physical fitness (Fühner et al., 2021). Additionally, age and maturity offset will also be considered on an individual level to analyse whether the positive correlation between age-related effects and overall physical fitness is caused/moderated by more mature children.

5 Analyses of maturity-related effects on physical fitness

5.1 Introduction

Finding effects of body height and body mass on physical fitness raises the question how growth-related changes are associated with different components of physical fitness. To elucidate these effects, this chapter analyses the effects of age and maturity offset (Mirwald et al., 2002; Moore et al., 2015) on physical fitness in children with deficits in their physical fitness.

5.2 Statistical analysis and model fitting

Effects of maturity offset, age, and gender were analysed for four physical fitness tests, namely 6 min run, star run, 20m sprint, and standing long jump (Fühner et al., 2021). Maturity offset was estimated using formulas by Mirwald et al. (2002) and Moore et al. (2015). Details of data pre-processing of physical fitness data can be found in Section 8.1.4 and of maturity offset data in Section 8.1.2. To avoid collinearity for age and maturity offset (since maturity offset is calculated using age, see Section 8.1.2), the first and second unrotated principal components were calculated based on z-scores of these variables (Hotelling, 1933) (for an application, see (R. M. Kliegl, 1982)) to achieve uncorrelated constructs (Section 8.4.1). The first principal component is calculated as mean of z-scores for age and maturity offset. This proxies relative growth, as higher values represent older children who are also closer to their individual growth spurt, while lower values represent younger children further away from it. The second principal component is calculated as difference between age and maturity offset z-scores, which represents delay in growth relative to chronological age. Positive values therefore represent children with greater skeletal growth delays, while negative values represent children with advanced skeletal growth.

The analyses involved two separate LMMs, one for each estimate of maturity offset. In both LMMs, physical fitness tests, relative growth, growth delay, and gender were included as fixed effects. Using LRTs and a parsimonious model selection approach (D. M. Bates et al., 2018), the best fitting LMMs including Test, relative growth, growth delay, and gender were selected and reported. Child was included as a random factor. Successive difference contrasts were set for the variables gender and Test. Residual-based diagnostics (e.g., q-q plot, standardized residuals over fitted values) were carried out for both final LMMs to assess model performance and are reported in Section 5.3.1.3 and Section 5.4.1.3. Means and standard deviations are reported in the Appendix Section 8.5.4.

The process of LMM selection for both models is documented in Section 8.5.2.1 and Section 8.5.3.1. The final LMMs included Test, relative growth, growth delay, and gender as fixed effects, with relative growth, growth delay, and gender specified as nested within the factor levels of Test. Child was included as a random factor with relative growth and growth delay further added as individual variance components.

5.3 Model 1: Effects of relative growth, growth delay, and gender on physical fitness (maturity offset estimate with Mirwald formula)

5.3.1 Results

5.3.1.1 Fixed effects

The LMM revealed a significant positive effect of relative growth on performance for star run ($\beta = .32$; SE = .06; z = 5.34; p < .001), 20 m sprint ($\beta = .13$; SE = .06; z = 2.21; p = .027), and standing long jump performances ($\beta = .24$; SE = .06; z = 4.03; p < .011). No significant effects were found for 6 min run performance ($\beta = -.11$; SE = .06; z = -1.8; p = .071), however, z-values indicate a negative trend.

Positive growth-delay effects were found for 6 min run ($\beta = .54$; SE = .13; z = 4.24; p < .001), 20 m sprint ($\beta = .38$; SE = .13; z = 3; p = .003), and standing long jump ($\beta = .33$; SE = .13; z = 2.6; p = .009), with better performances in children with a larger growth delay (i.e., smaller maturity offset score relative to their chronological age).

Gender-related differences were found for 6 min run ($\beta = -.75$; SE = .17; z = -4.49; p < .001), and 20 m sprint ($\beta = -.55$; SE = .17; z = -3.29; p = .001), with better performances in girls compared to boys.

Results, also including those for other tests, are reported in Table 5.1 and shown in Figure 5.1.

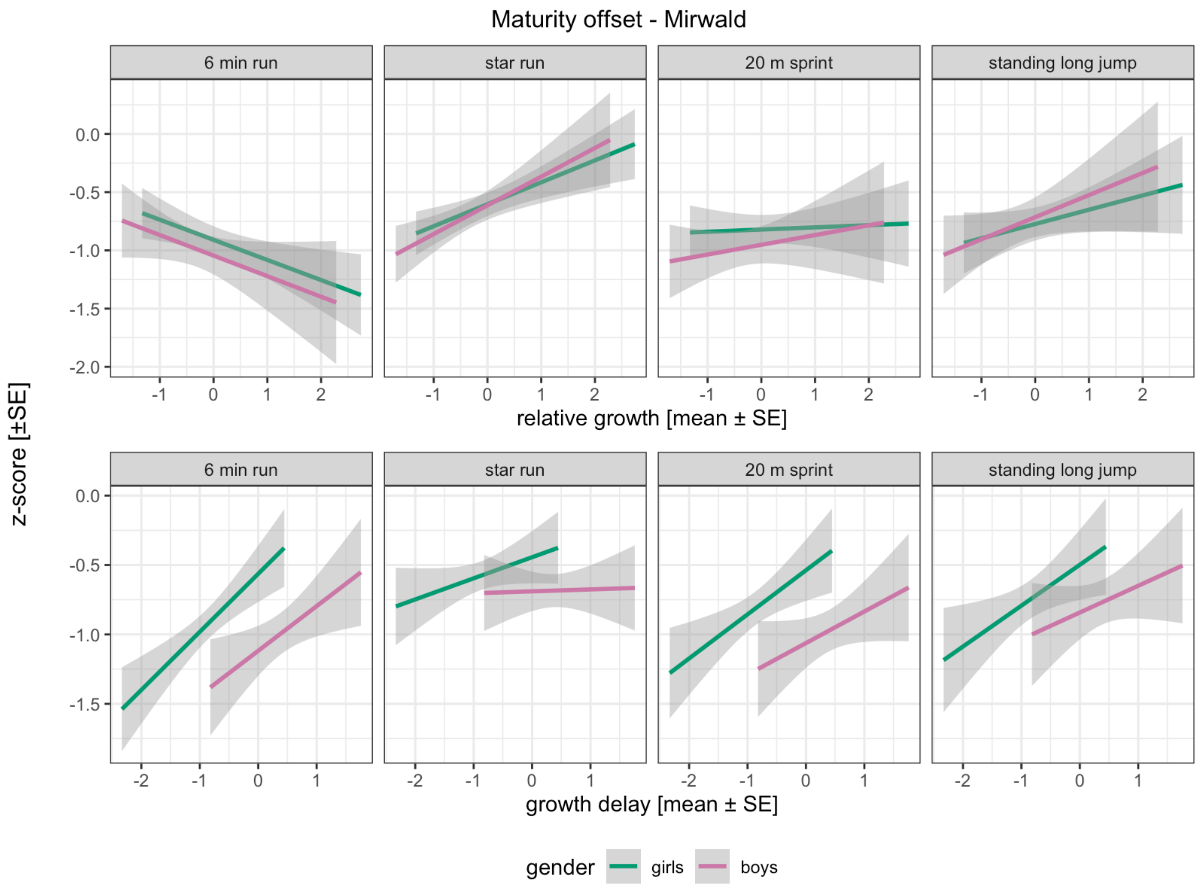


Figure 5. 1: Plots of overall effects for relative growth, growth delay, and gender [Mirwald]

Table 5. 1: Effects of relative growth, growth delay, and gender on physical performances [Mirwald]

Predictors	Est.	SE	z	p
Test [Run]	-0.81	0.08	-10.64	**<0.001**
Test [Star_r]	-0.54	0.08	-7.07	**<0.001**
Test [S20_r]	-0.76	0.08	-10.03	**<0.001**
Test [SLJ]	-0.63	0.08	-8.33	**<0.001**
Test [Run] : m1rg	-0.11	0.06	-1.80	0.071
Test [Star_r] : m1rg	0.32	0.06	5.34	**<0.001**
Test [S20_r] : m1rg	0.13	0.06	2.21	**0.027**
Test [SLJ] : m1rg	0.24	0.06	4.03	**<0.001**
Test [Run] : m1gd	0.54	0.13	4.24	**<0.001**
Test [Star_r] : m1gd	0.14	0.13	1.15	0.251
Test [S20_r] : m1gd	0.38	0.13	3.00	**0.003**
Test [SLJ] : m1gd	0.33	0.13	2.60	**0.009**
Test [Run] : gender2-1	-0.75	0.17	-4.49	**<0.001**
Test [Star_r] : gender2-1	-0.11	0.17	-0.69	0.493
Test [S20_r] : gender2-1	-0.55	0.17	-3.29	**0.001**
Test [SLJ] : gender2-1	-0.29	0.17	-1.76	0.079
Random Effects				
σ2	0.39			
τ00Child	0.19			
τ11Child.m1rg	0.08			
τ11Child.m1gd	0.31			
ρ01	-0.06			
	0.51			
ICC	0.33			
N Child	76			
Observations	1413			
Marginal R2 / Conditional R2	0.150 / 0.430			

Run = 6 min run, Star_r = star run, S20_r = 20 m sprint, SLJ = standing long jump, m1rg = relative growth (Mirwald) [z-score], m1gd = growth delay (Mirwald) [z-score], Est. = estimate; SE = standar error

5.3.1.2 Variance components and correlation parameters

Reliable individual variance components were found for relative-growth and growth-delay effects. A positive correlation was found between grand mean and growth delay effect (r = .51), showing children with better physical fitness also have a larger effect associated with growth delay. A negative correlation was found between relative-growth and growth-delay effects (r = -.74), revealing children with larger relative-growth effects tend to have smaller effects associated with growth delay. Results are reported in Table 5.2 and depicted in Figure 5.2.

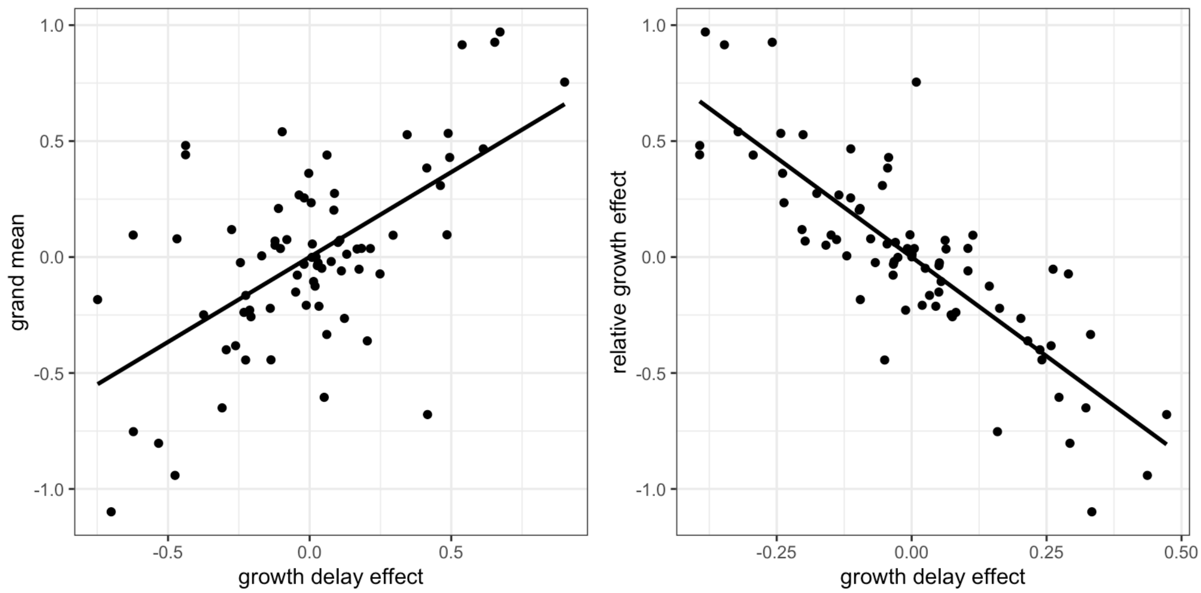


Figure 5. 2 : Correlations of grand mean and growth delay effect, and relative growth effect and growth delay effect in individual variance components [Mirwald]

Table 5. 2 : Variance components and correlation parameters [Mirwald]

Groups	Name	Variance	SD.	Corr	
Child	grand mean	0.191	0.437		
Child	m1rg	0.077	0.278	-0.062	
Child	m1gd	0.314	0.560	0.506	-0.752
Residual		0.390	0.624		

m1rg = relative growth (Mirwald), m1gd = growth delay (Mirwald)

5.3.1.3 Model diagnostics

Diagnostic checks of LMM residuals are plotted in Figure 5.2 and did not reveal any violations of LMM requirements.

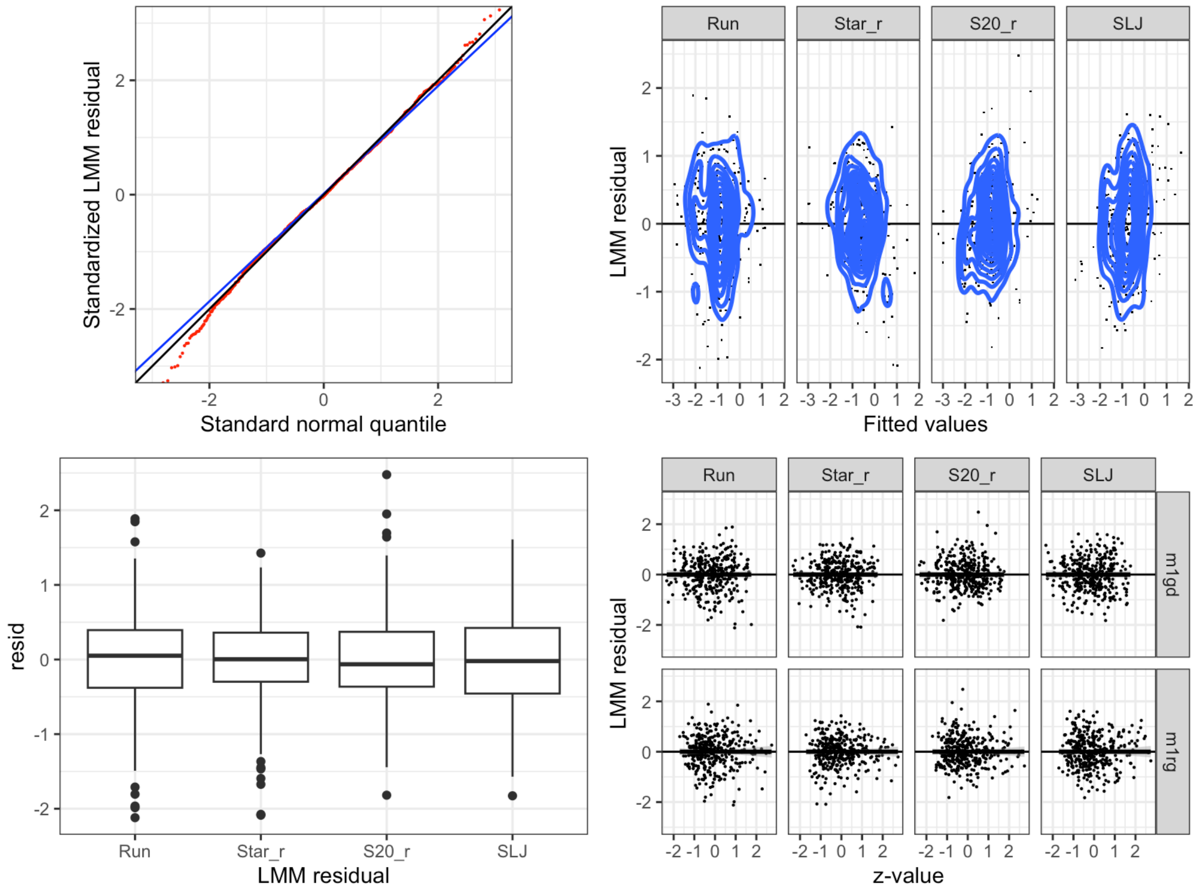


Figure 5. 3 : Model performance parameters

5.4 Model 2: Effects of relative growth, growth delay, and gender on physical fitness (maturity offset estimate with Moore formula)

5.4.1 Results

5.4.1.1 Fixed effects

The LMM revealed significant positive effects of relative growth for star run (β = .34; SE = .05; z = 6.48; p < .001), 20 m sprint (β = .2; SE = .05; z = 3.84; p < .001), and standing long jump (β = .3; SE = .05; z = 5.74; p < .001) performances.

For growth delay a positive significant effect was found on 6 min run (β = .44; SE = .18; z = 2.44; p = .015), with children having a larger growth delay eliciting better performances.

Further, a significant effect for gender was found in 6 min run (β = -.63; SE = .22; z = -2.88; p = .004), with girls outperforming boys.

Results are reported in Table 5.3 and shown in Figure 5.4.

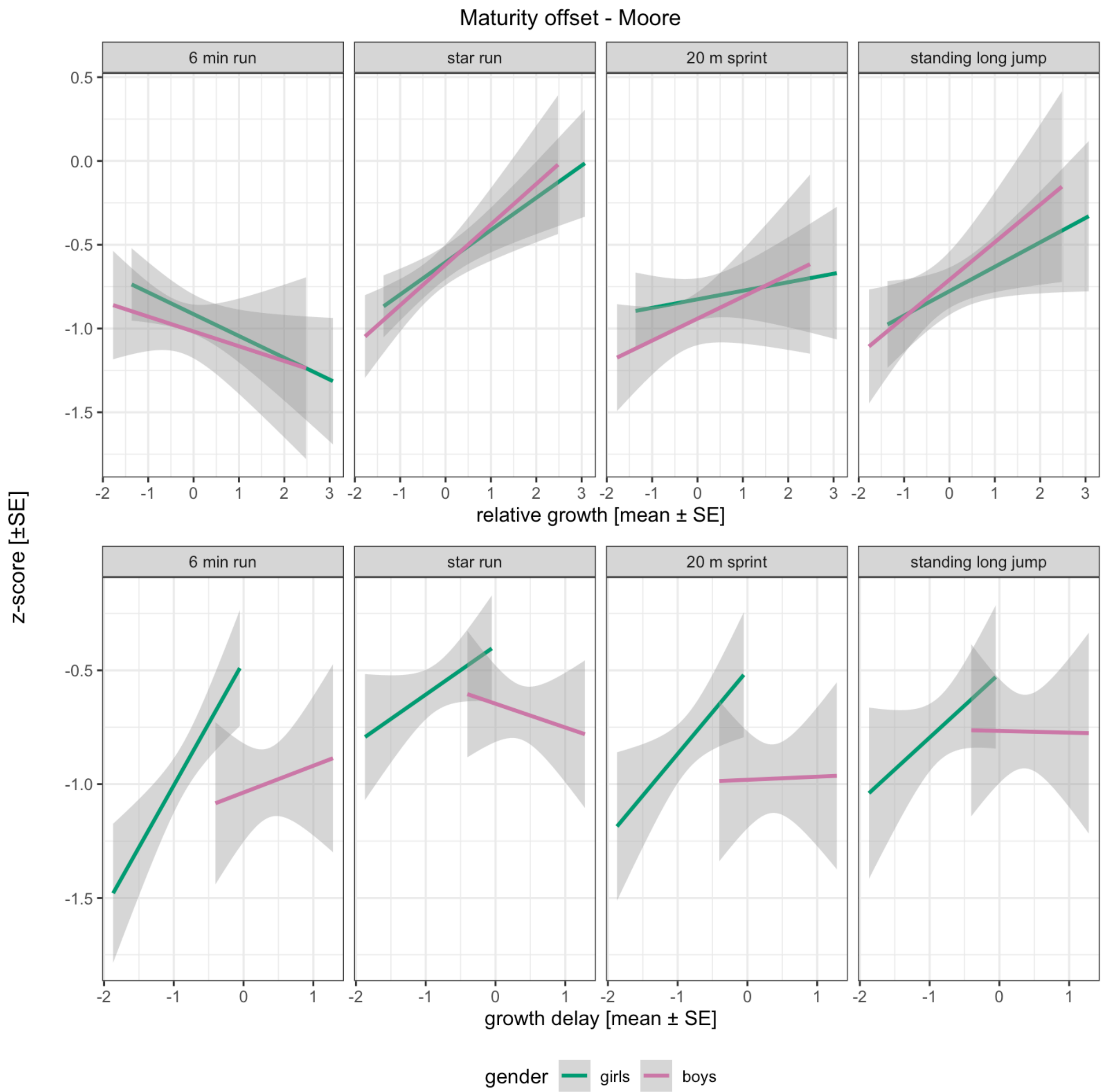


Figure 5. 4: Plots of overall effects of relative growth, growth delay, and gender [Moore]

Table 5. 3: Effects of relative growth, growth delay, and gender on physical performance [Moore]

Predictors	Est.	SE	z	p
Test [Run]	-0.82	0.09	-9.30	**<0.001**
Test [Star_r]	-0.56	0.09	-6.43	**<0.001**
Test [S20_r]	-0.79	0.09	-9.06	**<0.001**
Test [SLJ]	-0.67	0.09	-7.72	**<0.001**
Test [Run] : m2rg	-0.02	0.05	-0.36	0.722
Test [Star_r] : m2rg	0.34	0.05	6.48	**<0.001**
Test [S20_r] : m2rg	0.20	0.05	3.84	**<0.001**
Test [SLJ] : m2rg	0.30	0.05	5.74	**<0.001**

Table 5. 3: Effects of relative growth, growth delay, and gender on physical performance [Moore] (continued)

Predictors	Est.	SE	z	p
Test [Run] : m2gd	0.44	0.18	2.44	**0.015**
Test [Star_r] : m2gd	0.01	0.18	0.04	0.965
Test [S20_r] : m2gd	0.20	0.18	1.09	0.276
Test [SLJ] : m2gd	0.11	0.18	0.60	0.548
Test [Run] : gender2-1	-0.63	0.22	-2.88	**0.004**
Test [Star_r] : gender2-1	-0.00	0.22	-0.00	0.998
Test [S20_r] : gender2-1	-0.34	0.22	-1.58	0.115
Test [SLJ] : gender2-1	-0.05	0.22	-0.23	0.817
Random Effects				
σ2	0.40			
τ00Child	0.28			
τ11Child.m2rg	0.04			
τ11Child.m2gd	0.40			
ρ01	0.00			
	0.71			
ICC	0.42			
N Child	76			
Observations	1413			
Marginal R2 / Conditional R2	0.105 / 0.478			

Run = 6 min run, Star_r = star run, S20_r = 20 m sprint, SLJ = standing long jump, m2rg = relative growth (Moore) [z-score], m2gd = growth delay (Moore) [z-score], Est. = estimate; SE = standar error

5.4.1.2 Variance components and correlation parameters

Similar to the LMM reported above a positive correlation of variance components were found for the grand mean and the second principal component (r = .71), with children who have larger growth-delay effects also having better overall physical fitness, and a negative correlation was found between the first and second principal component (r = -.5), with children who have larger relative growth-related improvements exhibiting smaller effects of growth delay on physical fitness, and vice versa. Results are reported in Table 5.4 and depicted in Figure 5.5.

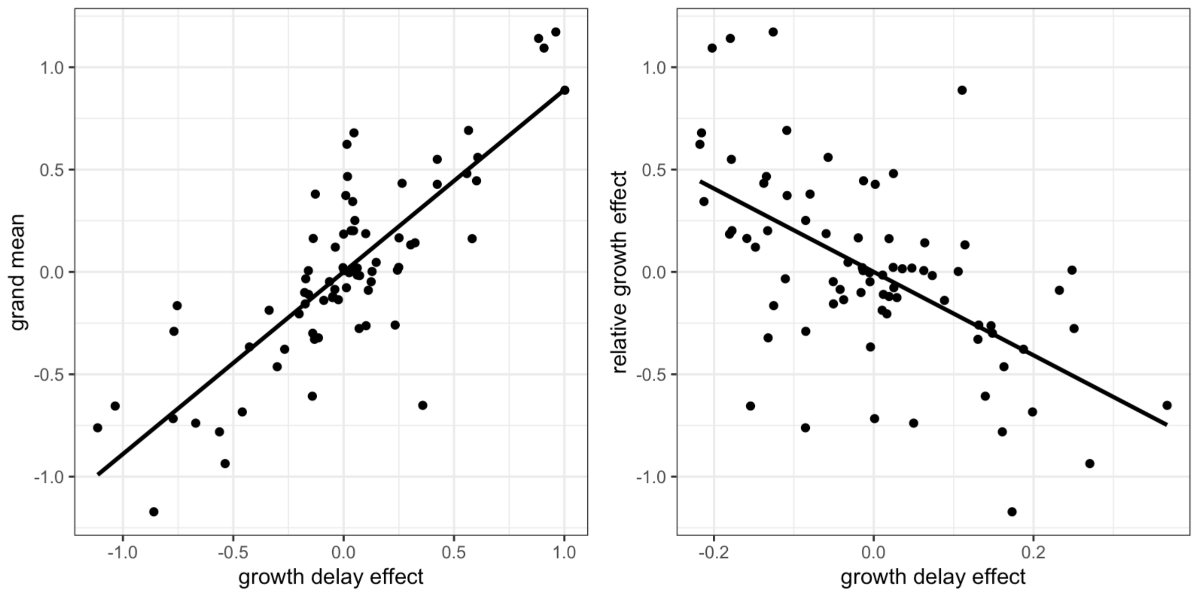


Figure 5. 5: Correlations of grand mean and growth delay effect, and relative growth effect and growth delay effect in individual variance components [Moore]

Table 5. 4: Variance components and correlation parameters [Moore]

Groups	Name	Variance	SD	Corr	
Child	grand mean	0.283	0.532		
Child	m2rg	0.042	0.204	0.001	
Child	m2gd	0.402	0.634	0.708	-0.501
Residual		0.396	0.629		
m2rg = relative growth (Moore), m2gd = growth delay (Moore)					

5.4.1.3 Model diagnostics

Diagnostic checks of residuals are plotted in Figure 5.5 and did not reveal any violations of LMM requirements.

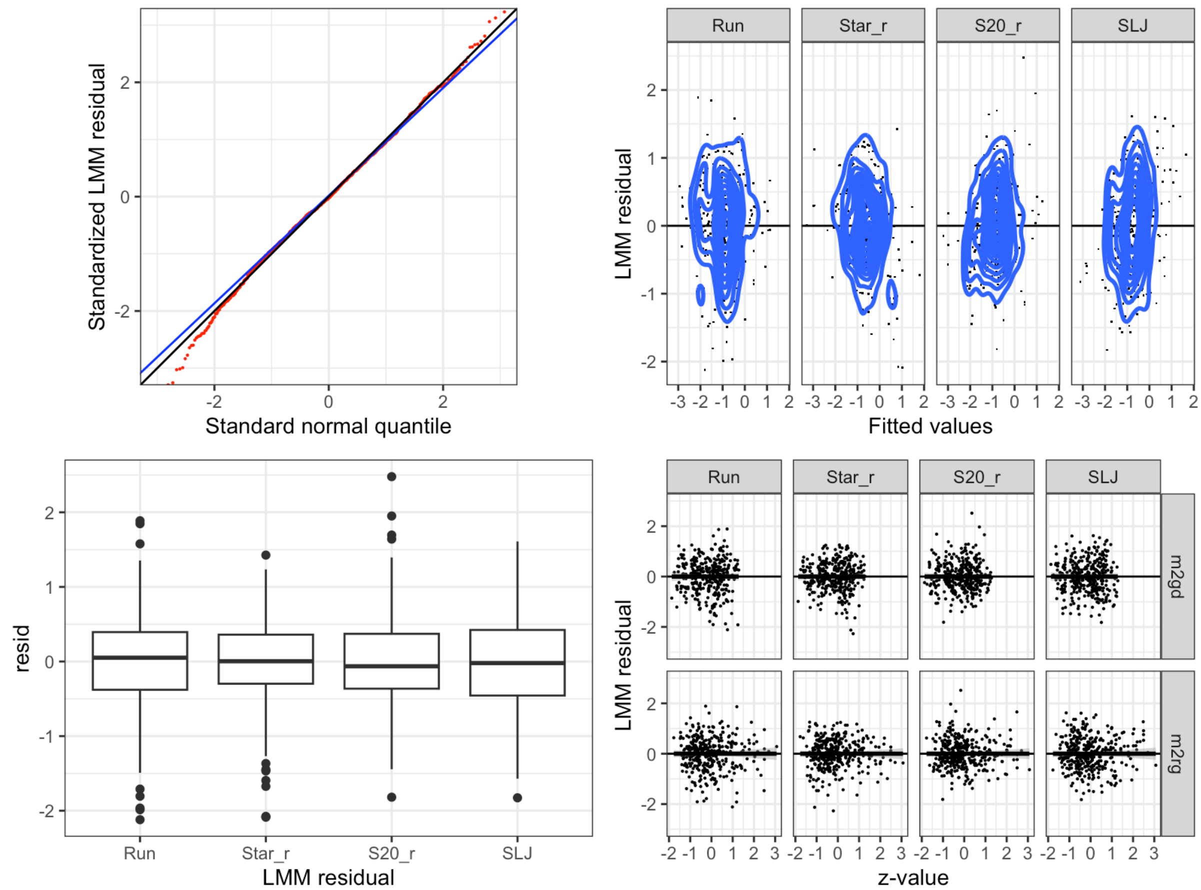


Figure 5. 6: Model performance parameters [Moore]

5.5 Summary of results

The main findings were (1) positive effects of relative growth for performances in star run, 20 m sprint, and standing long jump for both maturity offset formulas, with children of higher relative growth performing better, (2) positive effects of growth delay for performances in 6 min run for both maturity offset formulas, with children having larger growth delays showing better performances, (2.1) a positive effect of growth delay for performances in 20 m sprint and star run in the first LMM (i.e., Mirwald estimate) only, (3) gendered differences in 6 min run for both maturity offset formulas, with girls performing better than boys, and (3.1) gendered differences in 20 m sprint performances in the first LMM (i.e., Mirwald estimate) only, also with girls performing better than boys.

Both LMMs revealed relative growth and growth delay to be relevant individual variance components and further revealed (4) a positive correlation of individual variance components between overall physical fitness and growth delay, with children with greater growth-delay effects tending to have higher levels of physical fitness, and (5) a negative correlation of individual variance components between relative growth and growth delay, with children with larger relative growth-related effects tending to have smaller effects of growth delay on physical fitness and vice versa.

5.6 Discussion

To the author's knowledge, no comparable approach utilising unrotated principal components of age and maturity offset is available to this date. Further, comparable assessments investigating the development of physical fitness in children with physical fitness deficits is scarce, thus, generalisation of these results is limited. However, considering age-related improvements in physical fitness in regular fit children aged 8 to 14 years in these tests (Albrecht, 2015; Andersen et al., 1976; Castro-Piñero et al., 2009; Castro-Piñero et al., 2011; Catley & Tomkinson, 2013; Fühner et al., 2021, 2022; Golle et al., 2014; Lundgren et al., 2011; Miguel-Etayo et al., 2014; Niessner et al., 2020; Oliveira et al., 2014; R. Santos et al., 2014; Tambalis et al., 2016; Teich et al., 2023), the positive effect of relative growth for test performance in star run, 20 m sprint, and standing long jump to are not surprising and show that relative growth yields comparable results. Especially as age and maturity offset similarly progress over the consecutive assessments in the longitudinal design of the SMaRTER study, as well as the assessed components of physical fitness (as shown in Chapter 3). Interestingly, the model with maturity offset estimates by Mirwald et al. (2002) indicated a non-significant but potentially relevant trend for 6 min run performances, indicating a decline in performances in children of higher relative growth. The results would contradict the general assumption that age- and growth-related physical progress in terms of height, stature, and body composition leads to an improvement in physical fitness (Deprez et al., 2013; Lovell et al., 2015). While the plateau of cardiorespiratory fitness has been demonstrated in boys aged 15-16 years and in girls aged 14-15 years in physically fit children (Niessner et al., 2020; Ortega et al., 2011; Tambalis et al., 2016), the trajectory/onset might be different in children with deficits in physical fitness. Considering that maturity offset is a proxy for skeletal growth (Mirwald et al., 2002; Moore et al., 2015), a possible asynchronous progression of bone and muscle growth (e.g. bone growth at constant muscle capacity) (G. P. Beunen et al., 2006) would lead to a decrease in achieved force output (Hawkins & Metheny, 2001). As children in this study were within the range of growth spurt and peak elevation velocity (i.e., peak elevation velocity of -2 and 0 years) (Malina, 2004), growth-related changes might be more visible in continuous whole body workloads such as the 6 min run, where they would accumulate and thus explain negative non-significant effects of relative growth on 6 min run performances found in one model.

Effects for growth delay differed between models. The model using maturity offset according to Mirwald et al. (2002) found a positive growth delay effect on 6 min run, 20 m sprint, and star run performances, with larger growth delays leading to better performance. The model using the formula by Moore et al. (2015) found these effects on 6 min run only.

> **Side note:** Considering the novelty of this approach and variable, it might be worth elucidating growth delay using an example for the results for 6 min run. Older children (e.g., z-age = 1) that are further away from their age at peak height velocity (e.g., z-maturity-offset = -1) would yield higher growth-delay values (i.e., z-age - z-maturity offset = 1 - (-1) = 2), which would be indicative of a more delayed growth trajectory. In comparison younger children (e.g., z-age = -1) that are closer to their age at peak height velocity (e.g., z-maturity-offset = 1) would score lower growth-delay values (i.e., z-age - z-maturity offset = -1 - 1 = -2), which would be indicative of no delays or even an accelerated growth trajectory. Now regarding the positive effect found for growth delay on 6 min run performances, better performances would be found in children with larger growth delays, meaning for example same age children

with a delayed growth trajectory would tend to show better 6 min run performances compared to children with an accelerated growth trajectory.

This difference could be attributed to a difference in the complexity of formulas (see Section 8.1.2 for details). Moore et al. (2015) calculated the maturity offset based on age and standing or sitting height, while Mirwald et al. (2002) set up a more complex formula including multiple variables (i.e., standing height, sitting height, leg length, body mass, & age) while also giving age a less central role. Positive effects of growth delay on test performance could be explained by negative effects of skeletal growth discussed above. This means that children who enter the growth spurt later relative to their age experience fewer adverse effects of skeletal growth. One possible explanation could be that age, compared to skeletal growth, might be a better proxy for maturity- and other growth-related processes that improve fitness performance, such as muscle growth (G. P. Beunen et al., 2006). However, age-related improvements in physical fitness might also be influenced by social aspects related to age. Age, for example, regulates access to spaces that provide physical fitness services, such as school enrollment (Fühner et al., 2022) or access to organised sports facilities (Deprez et al., 2013; Lovell et al., 2015). Another example would be that all physical fitness recommendations in sport science include age-specific recommendations (Benjamin & Glow, 2003; Bull et al., 2020), which in turn are used to develop organised sport programs and curricula for physical education. This would mean that the environment might at least partially determine the magnitude of age-related effects on physical fitness (as argued in the previous chapter in the context of the positive correlation of individual levels of overall physical fitness and age-related improvements; see Section 4.5). Thus, age-related physical fitness improvements in children should be considered in context of growth processes, such as tissue growth, as well as the social environment available to children depending on their age.

LMMs also revealed gender differences, with girls performing better in 20 m sprint (in the model using Mirwald's estimate) and 6 min run (both models). Differences in outcome between both LMMs could be attributed to both maturity offset formulas comprising different gendered estimates, which presumably behave differently in the LMM approach. The direction of these findings contradicts current evidence for children, according to which boys tend to perform better than girls, as was found in children from the EMOTIKON sample (Fühner et al., 2021, 2022; Teich et al., 2023). A common argument for gendered physical fitness differences with better performance in girls compared to boys is that girls enter growth spurt and achieve age at peak height velocity two years earlier than boys (Malina, 2004) (see e.g. Fühner et al. (2022) analysis of older than key-age children in terms of 20 min sprint performance). This argument would fit the distribution of maturity offsets between girls and boys, as maturity offset is on average ~1 year closer to age at peak velocity for girls than for boys (i.e., Mirwald: girls = -1.6 ± 1 & boys = -2.7 ±.8 years; Moore: girls = -1.8 ± .9 & boy = -2.7 ± .7 years). However, this argument would contradict results for growth delay, where greater growth delay leads to better physical fitness performances, identifying the growth spurt and associated growth-related changes as detrimental to 6 min run performances. This means that girls' performances tend to stagnate or decline as they progress through growth spurt towards age at peak height velocity. Accordingly, another possible explanation could be that gendered differences in physical performance vary for children with deficits in physical performance. While there is ample evidence of physiological and social benefits for regularly fit boys compared to girls, such as better oxygen utilisation (Armstrong & Welsman, 2007; Tomkinson et al., 2018), higher muscle mass (Kanehisa et al., 1995; Malina, 2004), easier access to sport (Haywood & Getchell, 2014), and gender-specific self-perceptions and perceptions of sport through parents (Fredricks & Eccles, 2005), these factors may not affect children with deficits in physical fitness in a similar way. Accordingly, further research is needed to assess gender differences and their origins across the whole spectrum of physical fitness.

Finally, the inclusion of relative growth and growth delay as individual variance components allowed for an assessment of individual differences and their correlations. Considering age-related effects to be driven by anthropometrical growth-related changes (see Section 4.5), a positive correlation would have been expected for relative growth as well. Contradictory to this expectation, a negative correlation was found for overall physical fitness and growth-delay effects but no correlation for overall physical fitness and relative-growth effects. Additionally, a negative correlation between individual relative-growth and growth-delay effects was found, meaning that children with greater relative growth-related effects tend to show smaller growth-delay-related effects and vice versa. This could further underline potential adverse effects associated with skeletal growth discussed above and might reemphasize social components to drive positive effects associated with age on physical fitness discussed in Section 4.5 and Section 7.3. Further, considering only children with physical fitness deficits in their growth spurt before age at peak height velocity were included, these correlations might be specific to this period as well as the physical fitness level of the included children. Accordingly, further research is needed to explore the impact of different growth as well as maturity-related processes on development of physical fitness in children and their physical fitness levels.

5.7 Subsequent analyses

The following Chapter 6 is concerned with a different aspect of physical fitness. Utilising the multiple assessment for each child at each assessment in the Simon task (i.e., 60 data points each), the effects of physical fitness on executive function assessed using the Simon task were analysed. Implementing a LMM approach, the large number of measurements allowed to assess the effects of physical fitness on variance in the Simon task as well as accounting for individual differences in the effects of different physical fitness tests on Simon task reaction times.

6 Quasi-experimental effects of physical fitness on reaction time in a Simon task

6.1 Introduction

Due to the unique data assessment in the SMaRTER study, where physical fitness information of children physical fitness deficits were obtained alongside a large quantity of executive function measurements in the Simon task. This data structure allowed for a detailed analysis approach which assessed how the different components of physical fitness might moderate Simon task reaction time. Further, as multiple assessments were available for each child, an approach assessing individual differences in these moderations was feasible. Accordingly, this chapter is concerned with the effects of EMOTIKON fitness tests on Simon task reaction time, further elucidating individual differences in these effects.

6.2 Statistical analysis

The central variable of this analysis is reaction time in the Simon task, where side (i.e., right or left side of the screen), colour (i.e., red or blue stimuli), or congruence (i.e., stimuli were on the same or opposite side of the correct answer button) of the stimulus were manipulated. Performance in 6 min run, 20 m sprint, star run, standing long jump, ball push test, and one leg balance were used as covariates. In addition, age of participants and gender were included as covariates. For details on pre-processing of physical fitness data, see Section 8.1.4 and for age data, see Section 8.1.2. As far as reaction times are concerned, only correct responses in the range of 500 to 10000 ms were taken into account. A Boxcox distribution analysis of reaction times (Box & Cox, 1964) suggested a reciprocal transformation to response speed [speed = 1000/reaction time] to obtain normally distributed model residuals (see Figure 6.1). There was a strong general trend in speed over time for congruence, side, and colour (see Figure 6.2).

Model fitting followed a parsimonious approach (D. M. Bates et al., 2018) and occurred in two steps. In the first model, the main effects congruence, side, and colour were included as fixed effects together with the covariates age and gender. Reliable fixed effects were entered into the individual random effects structure to account for variance between individuals. After selecting a model based on this specification, 6 min run, star run, 20 m sprint, standing long jump, ball push test, and one leg balance were added to the fixed effects. As with the base model, the reliable fixed effects were entered into the individual random effects structure to elucidate individual differences in the association between executive function and physical fitness components. Contrasts for congruence, side, colour, and gender were set using the 'EffectsCoding'-command from the MixedModel package in Julia. Model diagnostics for the distribution of variance components in the random effect structure in the fitted models were based on qq and caterpillar-plots and are reported in the results.

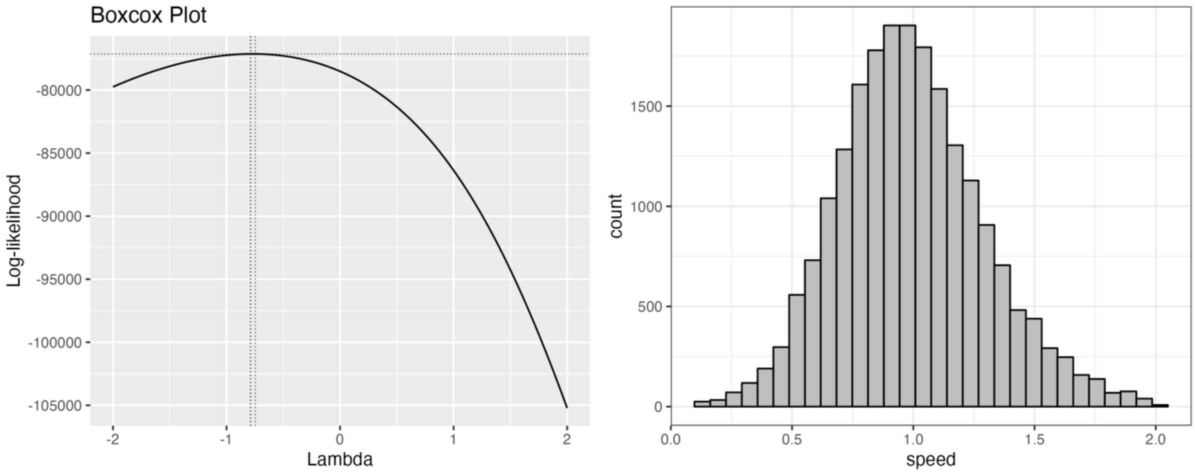


Figure 6. 1: Box-Cox distribution of reaction time and histograms of reaction speed

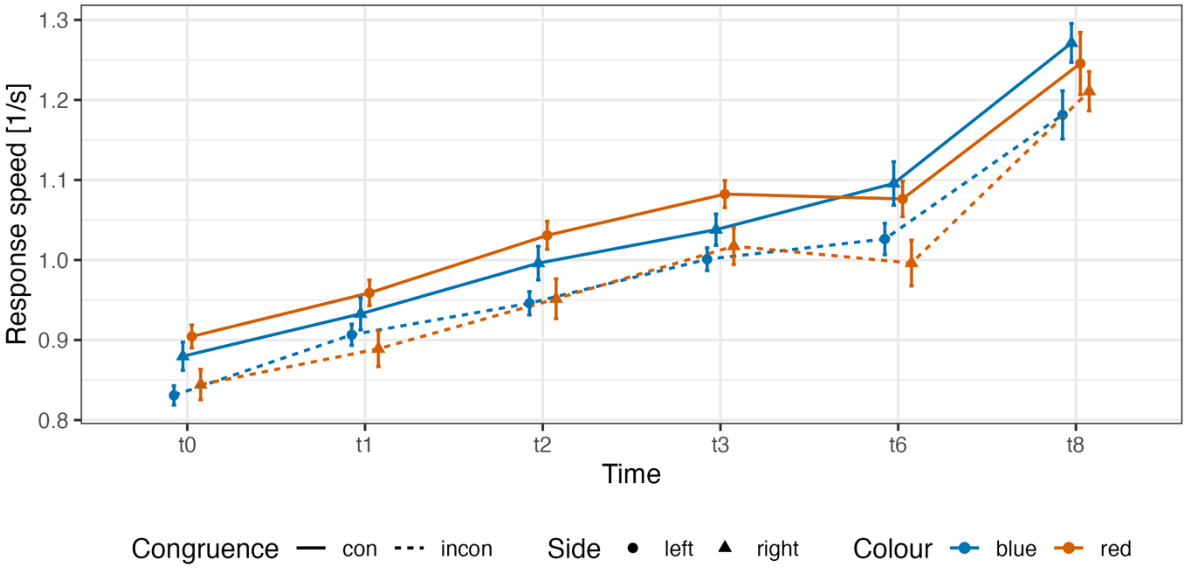


Figure 6. 2: Development of response speed for main effects in the Simon task across all assessments

6.3 Simon Task

6.3.1 Model fitting

Model selection is documented in Section 9.1.1. The final model was fitted with congruence, side, colour, age, and gender as fixed effects and variance components for grand mean, congruence, side, colour, and age in the random effect structure for each child. Correlation parameters were estimated for age and the three experimental effects of congruence, side, and colour.

6.3.2 Results

6.3.2.1 Fixed effects

The LMM for response speed in the Simon task revealed the expected effect of responding faster to congruent trials compared to incongruent ones (congruence: β = .029; SE = .003; z =

11.39; p < .001). Responses were also faster on the right than on the left side (side: β = -.005; SE = .002; z = -1.99; p = .046). The effect of colour was not significant. As for the covariates, reaction time increased significantly with age (age: β = .136; SE = .009; z = 15.59; p < .001) and boys showed a faster response speed compared to girls (gender: β = .04; SE = 0.016; z = 2.5; p = .012). The effects are shown in Figure 6.4 and reported in Table 6.1. Of note, depicting effects for side, across age and gender visualised a potential interaction for side and age. However, during the model selection process interactions between main effects with age and gender did not significantly improve the models fit and thus were not included in the final model. Means and standard deviation of included variables are displayed in Section 9.1.1.2.

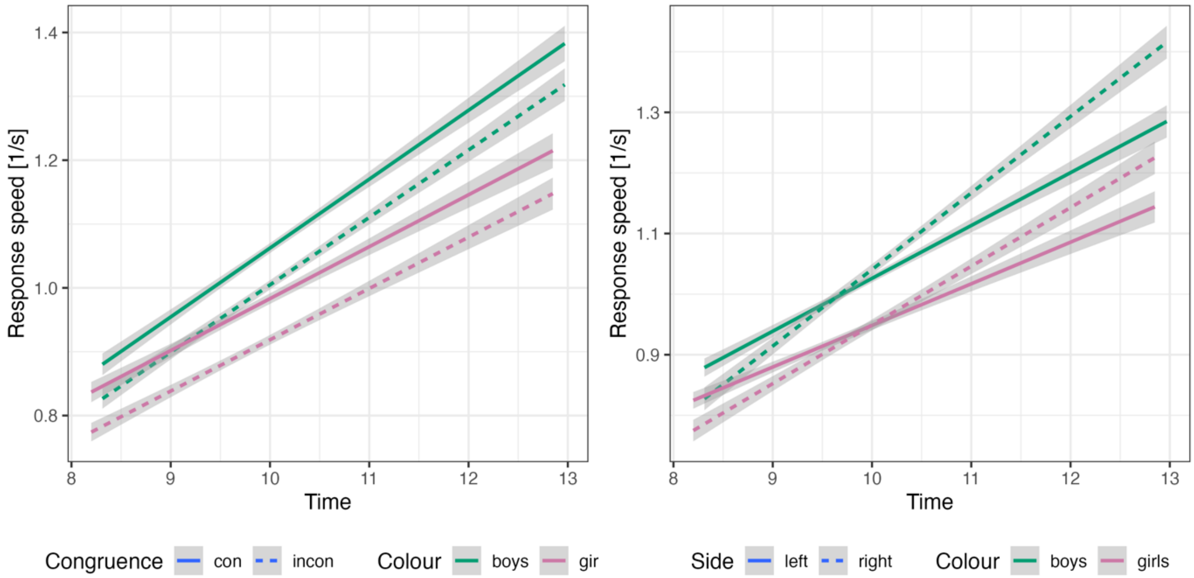


Figure 6. 3: Depiction of effects of Congruence, Side, age, and gender in Simon task reponse speed

Table 6. 1 : Fixed effects of the Simon Task main effects, age [a1], and gender

Name	Est.	SE	z	p
Grand mean	0.987	0.018	56.1	<0.001
Congruence: con	0.029	0.003	11.4	<0.001
Side: right	-0.005	0.002	-2.0	0.049
Colour: red	0.005	0.003	1.7	0.096
a1	0.136	0.009	15.6	<0.001
gender: boy	0.04	0.016	2.5	0.012

6.3.2.2 Variance components and correlation parameters

Reliable individual differences were found for the three experimental effects. Moreover, some of them correlated with each other. Specifically, effects of colour and congruence correlated with r = -.41, meaning that children who were more affected by congruence being less affected by different colours. Colour and side effects correlated with r = -.58, meaning that children who exhibited a larger Colour effect tended to exhibit a smaller Side effect on their reaction speed. Note that there are significant correlation parameters associated with the colour effect, although the fixed effect of colour was not significant. This can be explained by the fact that individual differences in the colour effect cancelled each other out in the overall effect.

Variance components and their correlations are reported in Table 6.2 and depicted in Figure 6.4.

Table 6. 2: Random effects of the Simon Task main effects and age [a1]

Groups	Name	Var	SD	Corr			
Child	Grand mean	0.15	0.023				
	a1	0.067	0.004	0.303			
	Congruence: con	0.016	0.0	0.019	0.14		
	Side: right	0.012	0.0	-0.093	0.197	-0.17	
	Colour: red	0.023	0.001	-0.386	0.132	-0.413	-0.576
Residual		0.24	0.058				

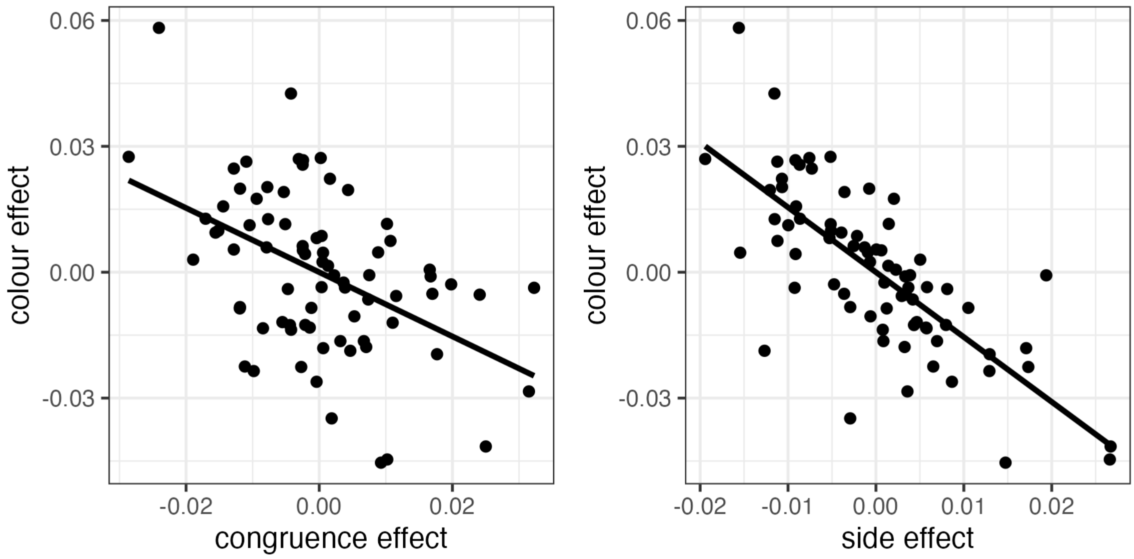
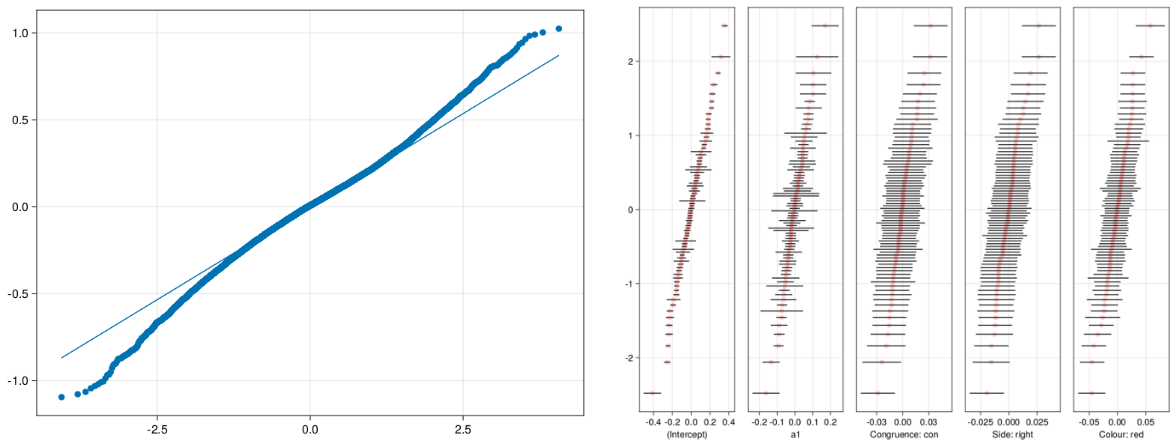


Figure 6. 4: Visual depiction of correlations of individual Simon task variance components

6.3.2.3 Model performance

A qq-plot and qq-caterpillar-plot (see Figure 6.4) of the LMM showed no outliers in the model residuals and residuals of the random effect structure of Child.



(a) qq-plot

(b) qq-caterpillar-plot

Figure 6. 5: Model diagnostics for m2

6.4 Simon Task x EMOTIKON

Using the previous LMM as a foundation, EMOTIKON Tests were included as fixed effects and as individual variance components.

6.4.1 Model fitting

The model selection process is documented in Section 9.1.2. The final LMM was fitted with Congruence, Side, Colour, age, gender, 6 min run, star run, and one leg balance test as fixed effects and with variance components for Grand Mean, Age, Congruence, Side, Colour, one leg balance, star run, and 6 min run in Child's random effects structure.

6.4.2 Results

6.4.2.1 Fixed effects

As reported above, the LMM revealed expected faster responses in the Simon task for congruent than incongruent trials (congruence: $\beta = .029$; $SE = .003$; $z = 11.42$; $p < .001$).

Effects of side and colour were not significant with EMOTIKON covariates in the model. As before, older children responded faster than younger children (age: $\beta = .136$; $SE = .01$; $z = 11.77$; $p < .001$).

The main interest was in effects of physical fitness test performances on response speed and their interactions with Simon task reaction speeds. None of these effects were significant (see Table 6.3).

However, we noticed that there was a negative association with

- 6 min run ($\beta = -.011$; $SE = .005$; $z = -2.37$; $p = .018$), and positive association with
- One leg balance ($\beta = .013$; $SE = .003$; $z = 4.67$; $p < .001$), and
- Star run ($\beta = .023$; $SE = .005$; $z = 4.74$; $p < .001$),

when the three tests were not included in the random effects for Child (see Section 9.1.2.1 for complete documentation). Means and standard deviation of included variables are displayed in Section 9.1.2.2.

Table 6. 3: Fixed effects of the Simon Task main effects, 6 min run [Run], star run [Star_r], one leg balance [OLB_l], age [a1], and gender

Name	Est.	SE	z	p
Grand mean	1.009	0.029	35.3	<0.001
Congruence: con	0.029	0.003	11.4	<0.001
Side: right	-0.004	0.002	-1.8	0.073
Colour: red	0.005	0.003	1.7	0.097
a1	0.136	0.012	11.8	<0.001
gender: boy	0.022	0.019	1.1	0.269
Run	0.004	0.02	0.2	0.86
Star_r	0.032	0.021	1.6	0.117
OLB_l	-0.004	0.014	-0.3	0.773

6.4.2.2 Variance components and correlation parameters

There were individual differences in the effects of age, Colour, Side, and Congruence, and for performance in the three EMOTIKON tests (6 min run, star run, and one leg balance) on response speed in the Simon task. Correlation parameter of the previous model were replicated (with slightly larger magnitudes, see Figure 6.4). There were also a relevant positive correlation between the grand mean and 6 min run effects (r = .45), a negative correlation between effects of age and star run (r = -.4), and negative correlation between effects of 6 min run and star run (r = -.46). As shown in Figure 6.6, children with a higher 6 min run effect had faster reaction times. Children with a higher star run effect on response speed had a smaller age effect, and children with a higher 6 min run effect tended to have a lower star run effect on response speed. Variance components and their correlations are reported in Table 6.4.

Table 6. 4: Random effects of the Simon Task main effects, 6 min run [Run], star run [Star_r], one leg balance [OLB_l], age [a1], and gender

Groups	Name	Var	SD	Corr						
Child	Grand mean	0.215	0.046							
	a1	0.077	0.006	0.08						
	Congruence: con	0.016	0.0	0.309	-0.112					
	Side: right	0.012	0.0	0.153	-0.407	0.147				
	Colour: red	0.024	0.001	-0.047	-0.661	-0.166	0.1			
	Run	0.147	0.022	-0.118	0.454	0.215	0.06	-0.226		
	Star_r	0.15	0.023	0.259	0.08	-0.398	-0.469	-0.107	-0.002	
	OLB_l	0.106	0.011	-0.071	0.23	0.073	-0.185	-0.051	-0.097	-0.16
Residual		0.054	0.054							

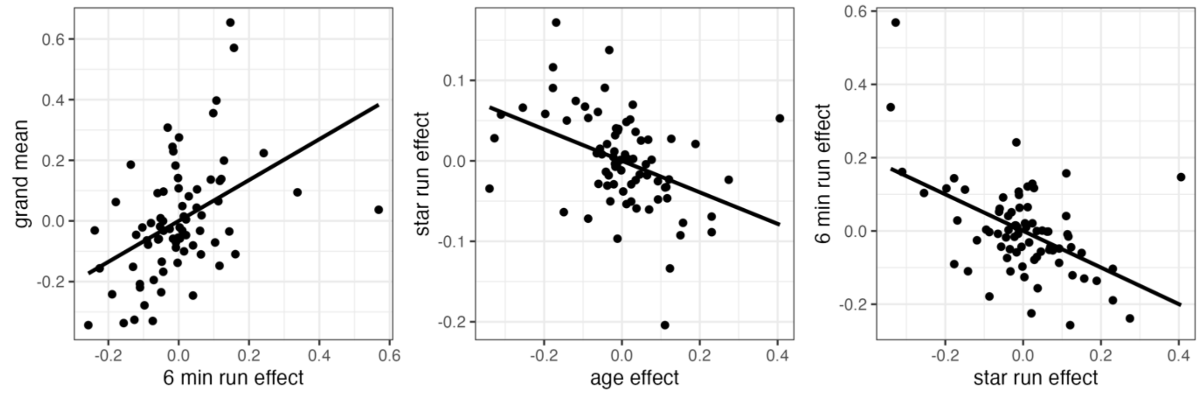
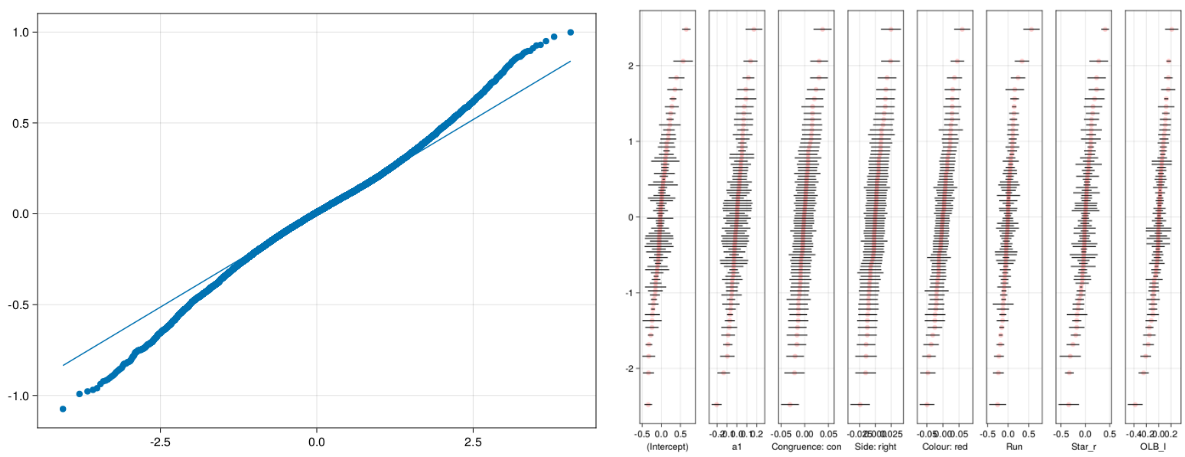


Figure 6. 6: Visual depiction of correlations of individual EMOTIOKON tests variance components

6.4.2.3 Model performance

A qq-plot and qq-caterpillar-plot (see Figure 6.7) of the LMM showed no outliers in the model residuals and residuals of the random effect structure of Child.



(a) qq-plot

(b) qq-caterpillar-plot

Figure 6. 7: Model diagnostics for m8

6.5 Summary of results

6.5.1 Fixed effects

Assessing the interrelations of physical fitness and executive function in children with deficits in their physical fitness, the main findings of both LMMs were (1) consistent congruence effects, with faster responses in congruent trials compared to incongruent trials, (2) consistent age and gender effects, with reaction speed increasing with age and being faster in boys compared to girls, and (3) a significant side effects, with faster response speeds on the left side compared to the right

The LMMs including the EMOTIKON tests revealed (4) significant positive effects for one leg balance and star run performances and a significant negative effect for 6 min run performances on Simon task response speed, when individual differences in these tests were not accounted for.

6.5.2 Variance components and correlation parameter

Both LMMs revealed the Simon task main effects to be relevant individual variance components and further revealed (5) a negative correlation for colour and congruence effects, indicating that children who were more affected by the congruent condition were less affected by different colours in their response speed, and (6) a negative correlation for colour and side effects, indicating that children who's Simon task speed was more affected by the side condition were less affected by different colours.

The LMM including EMOTIKON tests in the random effect structure of the children further showed 6 min run, star run, and one leg balance test to be relevant individual variance components and additionally showed (7) a positive correlation between grand mean and 6 min run effects, indicating children with a larger 6 min run effect had faster overall response speeds, (8) a negative correlation between star run and age effects, indicating that children with a larger star run effect had a smaller age effect, and (9) a negative correlation between 6 min run and star run effects, indicating that children with a larger 6 min run effect had a smaller star run effect.

6.6 Discussion

The significant effects for congruence, meaning faster response speeds for congruent compared to incongruent conditions, as well as overall improvements in response speed with age were consistent with current literature on Simon task performances in children (Davidson et al., 2006). A likely cause of congruence effects is the tendency to respond on the same side as the stimulus (Valle-Inclán, 1996), which is the dominant response that must be suppressed or inhibited to correctly identify incongruent stimuli (Bastian et al., 2016; Hilchey & Klein, 2011), causing a delay in response time. The improvement in reaction speed with age could be explained by growth- and maturity-related changes of physiological and psychological conditions that play a role in the Simon task, such as reaction to appearance of the stim- ulus with subsequent identification or execution of the respective motor response (Chan et al., 1991). Other aspects to consider are cultural and learning aspects associated with the task, such as increasing familiarity with a tablet, as well as learning from repeated test execution (S. C. Li et al., 2004).

Moreover, the first LMM showed effects for side and gender, with faster reaction times for boys compared to girls and for the left side compared to the right. As mentioned in Section 2.4.3, Simon task conditions were biased towards the left side (i.e., 40 left and 20 right), which may have biased children to pay more attention to the left side and thus achieving faster reaction speeds for stimuli appearing there. Another influence could result from general prevalences of right-handedness (Papadatou-Pastou et al., 2020) (handedness was not assessed in the SMaRTER study), as children may have covered parts of the right side of the screen with their right hand hovering above it in order to react quickly to presented stimuli. This could have shifted their attention more towards the left side. Interestingly, Figure 6.2 and Figure 6.3 show a possible interaction between side/colour and age, with a shift in side/colour occurring between t3 and t6, which frame the onset of the covid pandemic. However, since interactions were not included in the model, these results are not further discussed. In terms of gender, a study examining effects of different conditions of task switching using Simon task variations in 325 children aged 4 to 13 years found no gender differences for any conditions examined (Davidson et al., 2006). One possible reason for gendered differences found in this analysis might be a gendered upbringing where, for example, boys are more fixated on technical toys compared to girls (Todd et al., 2017). However, interpretation of gender differences in cognition must be done with caution, considering that the established field of gender neuroscience seems to be biased by the assumption of a "biological binary gender distribution" (R. M. Jordan-Young, 2010) and gender effects (as well as effects for side) did not remain significant when physical fitness components were entered into the model.

In both LMMs, correlations of individual variance components for colour and congruence effects as well as colour and side effects were negative. Since reaction speeds were faster for red as well as congruent stimuli, this correlation suggests that children who have slower reaction speeds in the incongruent condition (compared to the congruent condition) have faster reaction speeds for the blue stimulus (compared to the red stimulus). In other words, children who have greater cognitive flexibility (i.e., lower congruence effects) tend to have faster reaction times to red stimuli (i.e., larger colour effects). Studies have found faster reaction times for red stimuli compared to blue or grey stimuli across a range of cognitive assessments (Elliot, 2019; Elliot & Aarts, 2011), though they appear to vary depending on conditions like difficulty and type of cognitive task in which they were assessed (Elliot, 2019; Xia et al., 2016). Accordingly, since no significant main effect for colour was found in any of the models, this could imply that the colour effect might have been "overshadowed" or "swallowed" by the congruence effect. A similar dynamic could have caused the correlation of the individual variance components colour and side. However, this effect could also have been caused by the already mentioned skewness of the Simon task in favour of left stimuli.

Including 6 min run, star run, and one leg balance in the fixed effects revealed positive effects on reaction speed for star run and one leg balance performances and a negative effect of reaction speed for 6 min run performance. As star run and one leg balance are both tests placed in the coordinative domain of physical fitness, they are expected to involve cognitive aspect which might also have beneficial effects on reaction speed. Thus, the positive effect of star run and one leg balance performances on Simon task reaction speed are consistent with recent findings (Haverkamp et al., 2021; Marchetti et al., 2015; Niet et al., 2014). The negative effect of 6 min run performances on reaction speeds in the Simon task on the other hand were in the opposite direction than expected (Hillman et al., 2009; Kvalø et al., 2019; Niet et al., 2014). However, these fixed effects could not be replicated in a model in which individual variance of these tests was accounted for.

The inclusion of physical fitness tests as individual variance components revealed a positive correlation between overall reaction speed and 6 min run effects, showing that children who elicit greater 6 min run effects also tend to have faster overall reaction speeds. This could be explained by structural changes in the brain induced by higher levels of cardiorespiratory fitness associated with improvements in executive function (Chaddock et al., 2011, 2012; Chaddock-Heyman et al., 2014; Cotman et al., 2007, 2007; Hillman et al., 2008; Tyndall et al., 2018). It is also possible that faster reaction speed signifies enhanced executive functions that are associated with improved pacing, thereby linking better executive functions to improved 6 min run performance (Hyland-Monks et al., 2018). In addition, the direction could be moderated by the social environment, as environments that support physical activity or provide better access to increase cardiorespiratory fitness also provide better access to cognitive engagement (Brockman et al., 2009).

A negative correlation was found between star run and 6 min run effects on reaction speed, which means that children with a larger 6 min run effect are more likely to have a smaller star run effect and vice versa. This might suggest that there are different aspects associated with 6 min run or star run performances that promote executive function but might inhibit each other.

Even if a trade-off may be possible between the above-mentioned neurophysiological changes associated with either cardiorespiratory fitness or coordinative aspects that are present in star run but not in 6 min run (such as improved cerebellar activation and excitability (Diedrichsen et al., 2007)), it seems unlikely that these aspects have mutually exclusive structure in their effect on reaction speed. Other possible explanation could therefore be that, on one hand, physiological changes associated with high physical activity and thus better 6 min run have a positive effect on executive functions (Hillman et al., 2005, 2009). While on the other hand, children with low levels of physical activity might be more likely to be invested in sedentary activities such as computer games, which are beneficial for development of executive functions (Nuyens et al., 2019). This is then positively related to faster reaction times in the Simon task as well as improved star run performance, but may show no connection to 6 min run performances or its effect on reaction speed. Another aspect to consider at this point is that the negative correlation between 6 min run and star run effects could be caused by covid pandemic-related changes in daily life, as habitual physical activity likely decreased and screen time increased during the last two assessments (Nagata et al., 2020; S. C. E. Schmidt et al., 2020, 2022). Note that these effects were not examined due to large dropouts in both assessments during the pandemic (see Section 2.6).

Similarly, these interdependencies may also account for the negative correlation between individual variance components found for age and star run effects on reaction speed, meaning that children with greater age-related improvements in reaction speed tend to have a smaller star run-related improvements. It should be noted that both age and star run have a positive effect on individual reaction speed and accordingly influence each other in magnitude rather than direction. Here too, changes in daily life associated with the covid pandemic may have negatively influenced the magnitude of 6 min run effects through reduced physical activity, while increased screen time may have promoted age-related improvements in the Simon task linked to familiarity with computer interfaces, as well as improved executive functions (Kovacs et al., 2022). Considering the impact of lockdown measures in the analysis, inferences should also be made about SARS-CoV-2 infections, especially in light of accumulating evidence linking SARS-CoV-2 infections and post-covid symptoms with cognitive impairments such as cognitive deficits, executive dysfunction, increased risk of seizures, and others (Hall et al., 2022; Taquet et al., 2022; Yea et al., 2023). Although these more severe impairments are comparatively rare in infected children, they suggest that SARS-CoV-2 infections could affect cognitive development in children. The decline in comprehensive testing for SARS-CoV-2 infections (MBJS, 2022c, 2022d) combined with the higher prevalence of asymptomatic infections in children (Ravindra et al., 2022) allows speculation about possible SARS-CoV-2-related cognitive shifts in youth populations. However, these effects are still unclear today and need to be clarified in future research.

Overall, these results underscore the relationship between physical fitness and executive function in children, as well as considerations of individual differences in these relationships. However, since comparable information is scarce, the question arises whether these correlations only exist in children with physical fitness deficits or whether they can be extrapolated to physically fit children. More specifically, how different aspects of physical fitness (e.g., different components and their levels) are related to children's executive function, and how these relationships are associated with the lived environment and socioeconomic conditions.

Part III

Context and Conclusion

7 Socioeconomic context, Conclusion, and Outlook

7.1 Synopsis of results

The results of the analyses conducted in this thesis can be summarised as follows:

(1) No evidence was found for the effects of 14 weeks of additional physical education on physical fitness and executive functions in children with deficits in their physical fitness. The lack of detectable intervention-related effects might have been influenced by an insufficient intervention duration. In addition, interventions that are tailored to a specific component of physical fitness, rather than comprehensive exercises that encompass multiple components, might lead to detectable improvements in physical fitness, even in interventions with comparatively shorter durations. Finally, more insight is needed into how physical fitness behaves in children with physical fitness deficits, what the possible causes of these deficits are, and how these causes can be translated into meaningful interventions to improve children's physical fitness and health.

(2) Larger body height is associated with better physical fitness performances, while an increase in body mass has negative effects on performances in tests where body mass is used directly, but seems to be beneficial when it can be used as a counter mass. While body mass index (BMI) is predominantly used in exercise research as a proxy for overweight/obesity, the use of the composite score may conceal a more complex and interactive set of variables that influence physical fitness. As these appear to vary between the different components of physical fitness, these relationships should be explored in future research.

(3) Relative growth seems to have an overall positive effect on physical fitness, with the exception of cardiorespiratory fitness, where even negative effects are observed. Furthermore, delays in growth trajectories seem to be beneficial, especially for cardiorespiratory fitness. This could indicate possible adverse effects of skeletal growth that seem to be specific to certain components of physical fitness and appear to be more pronounced in children with lower relative-growth or age-related improvements in physical fitness. Therefore, further research is needed to clarify the role of age and growth-related effects on different components of physical fitness and their correlates.

(4) 6 min run, star run, and one-leg balance performances only show effects on reaction speed in the Simon task when individual differences are not accounted for. However, accounting for individual variance revealed that the effects of 6 min run and star run appear to be in polar domains affecting executive function. This could be explained by differences in daily life that either emphasise physical and cognitive engagement or promote cognitive but not physical engagement. However, further research is needed to elucidate the links between physical fitness and executive function, taking into account the children's social circumstances and daily life, as well as individual differences.

7.2 Discussion of the aims of the SMaRTER study

The central aim underlying the analyses carried out in this thesis was described in Section 2.1 as:

Analysis of the short-, mid-, and long-term intervention effects of a remedial physical education intervention in third grade children with deficits in their physical fitness.

Following the discussion on the lack of evidence for intervention-related effects in the SMaRTER study in Section 3.7, further studies on fitness component-specific and dose-dependent physical education-based fitness interventions are needed. These studies would allow a more informed approach to the development of a physical education curriculum for third and fourth grade children with physical fitness deficits, which was defined as the main objective of the SMaRTER study. However, there would be merit in reconsidering the assumption that children with physical fitness deficits have ”untapped potential” that can be realised through additional exercise. A comparable approach implemented in FITNESSGRAM® identified the physical fitness of children within a ”Healthy Fitness Zone”, ”In Need of Improvement”, or ”In Need of Improvement - Health Risk” based on cut-off values with a high probability of a successful identification of metabolic syndrome[101]. Implemented physical fitness tests focused on aerobic capacity[102],muscular fitness/mobility[103], and anthropometric information[104]. Based on test-specific cut-off values, the programme suggests that children are classified as ”fit” if they are in the ”Healthy Fitness Zone” in 5 out of 6 tests (or, without anthropometric information, in 4 out of 5 tests) (Cureton et al., 2013). In implementing this approach, a study investigating the relationship between physical fitness[105] and scores on two academic achievement tests[106] within two cohorts[107] in 2002/03 and 2007/08 identified four groups of children: Those who were (1) assessed as ”fit” in two consecutive years, (2) assessed as ”fit” in the first year only, (3) assessed as ”fit” in the second year only, and (4) not assessed as ”fit” in any year. Implementing this approach, a study assessing physical fitness[108] associations with scores in two academic performance tests[109] within two cohorts[110] in 2002/03 and 2007/08 identified four cluster of children: (1) score as ”fit” in two consecutive years, (2) score as ”fit” in the first year only, (3) score as ”fit” in the second year only, and (4) score as ”fit” in none of the years. They found that children who were not categorised as ”fit” in both years performed worse in both academic tests than the other three groups. They also found that socioeconomic factors moderated the correlation: Children from lower socioeconomic backgrounds, such as those eligible for free or reduced-price lunch, those from ethnic minority backgrounds, or those whose parents did not graduate from high school were more likely to score lower on both academic tests (London & Castrechini, 2011). Similariy, a study examining the relationship between socioeconomic status and reaching the ”Healthy Fitness Zone” found that boys[111] from more vulnerable socioeconomic positions[112] were more likely to fail the ”Healthy Fitness Zone” for curl-ups, while girls[113] from more vulnerable socioeconomic positions were more likely to fail several ”Healthy Fitness Zones” (i.e., body mass index, pacer labs, curl-ups, push-ups, and sit-and-reach) (Bohr et al., 2013). Another study that cross-sectionally examined the prevalence of 5,613,228 4th-12th grade students who met or did not meet several ”Healthy Fitness Zones” (i.e., PACER labs, push-up and curl-up test) from 2006/07 to 2016/17 and their moderation through different types of discrimination. They found that the proportion of children meeting ”Healthy Fitness Zones” increased overall over time, across all gender and racial/ethnic subgroups. However, the analysis revealed differences between gendered and/or racial/ethnic subgroups: Boys achieved more ”Healthy Fitness Zones” compared to girls, and white children achieved more ”Healthy Fitness Zones” compared to Hispanic and/or black children, with Hispanic and non-Hispanic black girls achieving the lowest scores. Further, they could show the differences between gender and racial/ethnic subgroups widening over time between cohorts. In addition, their analysis of home neighbourhood socioeconomic status[114] revealed that the proportion of children who do not reach the Healthy Fitness Zones increased as the percentage of households below the poverty line in the neighbourhood increased (Konty et al., 2020).

At this point, it is important to consider that the associations between socioeconomic status and physical fitness deficits are not solely momentary assessments. Living and growing up in poverty has been shown to have cumulative effects that moderate children’s morbidity and cognitive and socio-emotional development (Evans, 2003). For example, a study analysing the associations between body mass index and a cumulative poverty risk factor[115] in 329 adolescents born between 1983 and 1990 and interviewed at the ages of 9, 13 and 17 revealed a positive association between body mass index and cumulative poverty risk score, with children who spent their entire lives in poverty having a higher risk of being overweight (Wells et al., 2010). Incorporating these associations into the scientific foundation of the research question would reframe the understanding of the 'untapped potential' hypothesis. To improve physical fitness in children with physical fitness deficits, interventions would need to consider current socioeconomic individual and community positions and any potentially related moderators of physical fitness as well as their cumulative characteristics during childhood development to date. Accordingly, the next chapter reviews and discusses the available socioeconomic information from the SMaRTER sample and the evidence for associations between physical fitness and socioeconomic moderators.

7.3 Discussion of socioeconomic context

In Section 2.5.2 socioeconomic information of the children was described and showed children where income information were available (i.e., 46 out of 76), to be predominantly placed in households having below-average income per capita. It should be noted, socioeconomic information was not part of the inclusion criteria, it is, however, striking that filtering for physical fitness deficits shows a strong bias towards below-average income. This bias agrees with findings of Tomkinson et al. (2019) that show international secular trends in shuttle run performance to be positively correlated with progression of the Gini index towards income equality (see Section 1.1.3.1 for details). Assuming that financial costs of a healthy lifestyle are based on the population average income of a country or region, and considering the Gini index as a measure of overall societal participation in such an active lifestyle, children's physical fitness (or at least certain subcomponents such as cardiorespiratory fitness) could be strongly related to family income. This relationship is further strengthened by interrelation between physical fitness and weight status, which also moves along socioeconomic lines (see Section 1.2) and which also showed positive but weak effects of physical fitness interventions (see Section 1.4.2). Taking these associations into account would allow for a more contextualised evaluation of previously discussed results.

This, for example, allows for a different contextualisation of the lack of intervention effects in the SMaRTER study (see Chapter 3), coupled with overall small and infrequent evidence for improvements of cognition and physical fitness associated with qualitative school-based physical fitness interventions at post assessments (Garciá-Hermoso et al., 2021; García-Hermoso et al., 2020) and at follow-ups (Jurak et al., 2013; Lai et al., 2014). If children's physical fitness deficits are caused or at least moderated by low income or lack of resources (as discussed in the chapter above), then physical education or school-based interventions that do not address or change socioeconomic conditions of children's families would always fall short of making meaningful changes in children's health status. Or to put it another way: Physical fitness interventions that are divorced from social change approaches to eradicate poverty only target the problem at the symptom level and will not be able change the overall trajectory of public (child) health.

Effects of maturation and growth delay on physical fitness explored in Chapter 4 may be related to theories suggesting that psychosocial stress in a family due to insufficient financial resources or lack of father figures leads to a "faster reproductive strategy" that causes earlier maturation/puberty (Hochberg & Belsky, 2013; Oelkers et al., 2021). Accordingly, in addition to limited access to resources to promote physical fitness, the socioeconomic environment would potentially influence the development of physical fitness by driving maturation/puberty. These advances could also affect skeletal growth, which in turn affects the development of physical fitness of children from lower income families differently than children from higher income families.

With the positive correlation between children's overall physical fitness and age-related effects found in Chapter 5 and the correlation between growth delay effects, maturation effects, and overall physical fitness in Chapter 4, moderating effects of socioeconomic background of families (Ma, 2000) and areas (Nieuwenhuis & Xu, 2021) were postulated. This would suggest that these associations are not only stable between families and areas across the whole socioeconomic spectrum, but are also present among children with low physical fitness from low-income families. Therefore, further research is needed to explore the impact of socioeconomic resources in families, areas, and schools in order to design meaningful interventions.

This would also be consistent with results of the Simon task in Chapter 6, where a positive correlation of individual variance components was found between overall reaction speed in the Simon task and 6 min run effects. Due to social mechanisms such as gentrification in the Berlin area (Borck & Gohl, 2021), higher income families tend to live in higher income neighbourhoods, which results in unequal distributions of resources with more resources available in more affluent areas and schools. Accordingly, they can offer their children a more physically and cognitively engaging environment, which may explain the EMOTIKON findings that children living near Berlin have better physical fitness compared to children living further away (R. Kliegl & Teich, 2022).

In addition, it should be noted that the impact of the covid pandemic was exacerbated in low-income families and areas which faced higher rates of infection and mortality (Elgar et al., 2020; Marmot & Allen, 2020), were associated with higher rates of pre-existing conditions (Heisig, 2021), had higher rates of job losses and income reductions and limited access to food and safe shelter (Wright et al., 2020), showed higher rates of financial stress (Ettman et al., 2021), had higher rates of depression and anxiety during lockdowns (Fancourt et al., 2020), and higher rates of parenting stress and mental health problems (J. Li et al., 2022), to name a few. However detrimental effects of the covid pandemic on physical fitness seem to be larger in higher income areas. For example, Teich et al. (2023) found larger declines in "fitter" schools compared to "unfitter" schools. They argue that these effects are due to the unequal distribution of physical fitness resources before the pandemic. While disadvantaged children have fewer resources to mitigate the loss of structured physical activities, they also had less access to these activities before the pandemic. Thus, larger declines can be found in more affluent areas/schools as there is "more to loose".

7.4 Integrative perspective

These findings allow to revisit the central question of the SMaRTER study: How to improve physical fitness in children who are the most at risk to develop adverse health effects? While literature assessing the effects of schoolbased interventions in the context of available family and community resources, such as family income (Ma, 2000; Tomkinson et al., 2019), or wealthy neighbourhoods/areas (R. Kliegl & Teich, 2022; Nieuwenhuis & Xu, 2021), is scarce, it seems evident that these parameters are central to children's physical fitness trajectories. This would raise the question, how can school-based physical fitness interventions be structured to effectively improve physical fitness and by extension population health and quality of life? However, poverty is a necessary condition for capitalist and colonial mechanisms of resource distribution (Marx, 1955; Sullivan & Hickel, 2023), which use systems of oppression such as racism and sexism to maintain social hierarchies (Mendívil & Sarbo, 2022). These systems of power create specific socioeconomic and cultural conditions that shape children's and families' environments and thus affect their health outcomes and life trajectories (see (Drewnowski & Eichelsdoerfer, 2010) for an analysis of the impact of poverty on obesity). Therefore, in order to develop meaningful interventions and policies need to take into account underlying socioeconomic conditions in order to improve public health and physical fitness.

Further, as shown in Chapter 4 and Chapter 5, anthropometrical parameters of children as well as their progression through age and growth-related development play an essential role in the development of physical fitness. While these parameters individually have been rigorously studied, interactions between different growth-related parameters and their effects on various physical fitness components is not well researched to this date. Accordingly, more research endeavours developing methods to elucidate these interactions as well as assessing their effects on physical fitness components in children are needed. In accordance with the paragraph above, these effects need to be properly contextualised in their environmental, cultural, and socioeconomical circumstances.

Finally, as shown in Chapter 6, multiple data points of the Simon task at any assessment allowed for a more detailed analysis and understanding of the interrelations of physical fitness and executive function through considering individual differences and variance. Accordingly, for future research endeavours assessing the interrelations of physical fitness and executive function and by extension any other potential correlates, individual differences and variance should be considered and integrated.

7.5 Conclusion

In conclusion, for the development of meaningful interventions of physical fitness, it is important to understand the relevance of the individual as well as environmental context that these interventions happen in. For the individual context, for example, the progression in growth and age and their interactions have shown to be relevant for developments of physical fitness. Regarding the environmental context, available resources such as income or time seem to be relevant causal moderators of physical fitness. These factors need to be considered in any assessment and integrated in interventions in order to achieve meaningful improvements in childrens physical fitness and general health.

Part IV

References and Footnotes

Notes

[1] Kraus et al. (1954) was criticized for alleged weaknesses in their research design and too few fitness components (i.e., flexibility and strength) (R. R. Pate, 1983).

[2] Relative leanness refers to the relation of lean body mass (i.e., weight of all body tissue except fat) to total body fat (Bubb, 1986).

[3] Including external stress as a determining factor allows to consider individual physical fitness and health in their socioeconomic conditions. Given associations of systematic social conditions with physical fitness (see e.g., Section 1.1.3.1 and Section 1.1.3.2) and weight status (see Section 1.2), this approach would allow these conditions to be taken into account when conceptualising research interests and developing policy guidelines.

[4] For a more detailed description and discussion of the approach of Bös et al. (2017) see Albrecht (Albrecht, 2015).

[5] Due to this assumption, summaries of studies include information about implemented tests, either in text or in footnotes, to clarify respective framings of physical fitness.

[6] Processes towards a mature state can be defined as growing of breasts, pubic hair, and age of menarche in girls (Marshall & Tanner, 1969), and growing of genitalia and pubic hair in boys (Marshall & Tanner, 1970).

[7] Another common approach is to determine sexual maturity using five stages of maturity (so-called Tanner stages) by screening and classifying growth of breasts and pubic hair in girls and genitalia and pubic hair in boys (Marshall & Tanner, 1969, 1970; Tanner & Whitehouse, 1981). These have been heavily criticised for their invasive/intrusive nature and gender bias (G. P. Beunen et al., 2006; Cameron, 2012; Mirwald et al., 2002). For an account of ethical concerns with data collection (e.g., consent) and historical context of Tanner's work on classifying maturity stages in children, see Roberts et al. (2016).

[8] The number of publications in which the terms 'sex' and 'gender' are implemented is increasing. However, there is a lack of clear definitions and concepts for these terms. For example, Tomkinson et al. (2019) examined international secular trends in cardiorespiratory fitness and reported on 'sex' specific trends and interactions. Their discussion then explored these differences in context of "gender equality policies and programmes". However, in their publication they do not address differences in these concepts and their implications.

[9] In their book "*Atlas of Children's Growth - Normal Variation and Growth Disorders*", Tanner and Whitehouse (Tanner & Whitehouse, 1981) classify children maturity according to their criteria, whereby children who can be easily classified according to their criteria are called "normal". Deviations from criteria are classified as "disorders" (as the title suggests). Some of their "disorder" classifications are, for example, variations in chromosome composition XO (Turner syndrome), XXY (Klinefelter syndrome) or XXY (XYY syndrome). For an account of stigmatisation of 'disorders' and processes of sex determination, see Meyer-Bahlburg et al. (2016) and Lee et al. (2016).

[10] Dihydrotestosterone induces formation of external sex organs and is crucial for the development of the infantile penis (Sobel et al., 2004).

[11] Attribution in this context occurs within frameworks of a binary understanding and in absence of first-hand information about gender identity of the other.

[12] Additional information may be required, for example, from athletes, such as assessment of hormone levels or determination of chromosomal structures. Here, identification can be alongside or in opposition to the others identity, and in most legal systems a binary understanding of gender and sex is still upheld and enforced (R. Jordan-Young & Karkazis, 2019).

[13] For an intersectional and postcolonial analysis of recent discussions on hyperandrogenism among professional female athletes, see (Karkazis & Jordan-Young, 2018).

[14] In line with Lugones (Lugones, 2010), this placement can be within and outside the binary 'sex/gender/sexuality system' (Seidman, 2009; Westbrook & Schilt, 2014).

[15] Cardiorespiratory fitness was assessed in cross-sectional studies using 20-m shuttle run in 1,026,077 children from 30 European countries (Austria, Belgium, Bosnia and Herzegovina, Croatia, Cyprus, Czech Republic, Denmark, Estonia, Finland, France, Germany, Greece, Hungary, Ireland, Iceland, Italy, Kosovo, Latvia, Lithuania, North Macedonia, Norway, Poland, Portugal, Slovakia, Spain, Serbia, Slovenia, Sweden, Switzerland) aged 6 - 18.9 years from ~2000 to 2023 (Ortega et al., 2023), 20-m shuttle run in 22,048 Portuguese children and adolescents (11,373 girls; 10-18 years) in 2008 (R. Santos et al., 2014), 20-m shuttle run among 424,328 Greek

children and adolescents (49% girls; 6-18 years) in 2014 (Tambalis et al., 2016), 20-m shuttle run, 1 mile, 1/2 mile and 1/4 mile run/walk in 2,752 Spanish children and adolescents (1,261 girls; 6-17.9 years) (Castro-Piñero et al., 2011), physical work capacity 170 static bicycle ergometer test in 3,742 German children and adolescents (50.1% girls; 4-23 years) from 2009 to 2013 (Niessner et al., 2020), 20-m shuttle run in 6,398 children from Sweden, Germany, Hungary, Italy, Cyprus, Spain, Belgium and Estonia (3,288 girls; 6-10.9 years) from 2007 to 2008 (Miguel-Etayo et al., 2014), 1-mile run/walk in 3,804 Portuguese children (1,819 girls; 6-10 years) (Oliveira et al., 2014), 6-minute run in 108,295 German children (8-9 years) (Fühner et al., 2021), and a meta-analytic approach involving 15 studies of a 1.6-km run or a 20-m shuttle run in Australian children (5-18 years) from 1985 to 2009 (Catley & Tomkinson, 2013). Longitudinal studies assessed cardiorespiratory fitness using 9-minute run in 240 German children (88 girls) aged 9-12 years in four consecutive annual assessments (Golle et al., 2014), assessment of maximal aerobic power in 65 children (34 girls) aged 8.4 to 12.4 years in five consecutive annual studies from 1969 to 1973 (Andersen et al., 1976), and also analysed in a review of longitudinal studies from Australia (3), Belgium (4), Canada (3), China (1), Denmark (4), Germany (10), England (3), Estonia (2), France (1), Greece (1), Japan (1), Luxembourg (1), Netherlands (1), New Zealand (2), Northern Ireland (1), Norway (1), Occupied Palestine (1), Portugal (1), Sweden (3), Switzerland (1), South Africa (1), and USA (1) (Albrecht, 2015).

[16] Muscular fitness was assessed in cross-sectional studies using hand grip strength in 747,966 children and standing long jump in 1,345,159 children from 34 European countries (Albania, Austria, Belgium, Bosnia and Herzegovina, Bulgaria, Croatia, Cyprus, Czech Republic, Denmark, Estonia, Finland, France, Germany, Greece, Hungary, Ireland, Iceland, Italy, Kosovo, Latvia, Lithuania, Luxembourg, North Macedonia, Netherlands, Norway, Poland, Portugal, Slovakia, Spain, Serbia, Slovenia, Sweden, Switzerland) aged 6 - 18.9 years from ~2000 to 2023 (Ortega et al., 2023), ball throw, standing long jump, vertical jump, push-up, hanging with bent arm, pull-up, sit-up, curl ups in 30 seconds and curl up tests in 2,778 Spanish children and adolescents (1,265 girls; 6-17.9 years) (Castro-Piñero et al., 2009), push-up and curl-up tests in 22,048 Portuguese children and adolescents (11,373 girls; 10-18 years) in 2008 (R. Santos et al., 2014), isokinetic peak concentric torque at 60 and 180⁄sec of knee extensors and flexors and vertical jump height in 436 Swedish children (190 girls; 6-12 years) from 1999 to 2001 (Lundgren et al., 2011), standing long jump and sit-up test in 424,328 Greek children and adolescents (49% girls; 6-18 years) in 2014 (Tambalis et al., 2016), handgrip strength and standing long jump in 3,804 Portuguese children (1,819 girls; 6-10 years) (Oliveira et al., 2014), sit-up and push-up tests and standing long jump in 3,742 German children and adolescents (50.1 % girls; 4-23 years) from 2009 to 2013 (Niessner et al., 2020), handgrip strength and standing long jump in 8,418 and 8,494 children from Sweden, Germany, Hungary, Italy, Cyprus, Spain, Belgium and Estonia (4,304 and 4,339 girls, respectively; 6-10.9 years) from 2007 to 2008 (Miguel-Etayo et al., 2014), ball push test and standing long jump in 108,295 German children (8-9 years) (Fühner et al., 2021), and another meta-analytic approach with 15 studies examining basketball throw, push-up test, standing long jump, sit-up test and handgrip strength in Australian children (5-18 years) from 1985 to 2009 (Catley & Tomkinson, 2013). Longitudinal studies assessed cardiorespiratory fitness using 1-kg ball throw test and triple hop test in 240 German children (88 girls) aged 9-12 years in four consecutive annual assessments (Golle et al., 2014), and also analysed in a review of longitudinal studies from Australia (3), Belgium (4), Canada (3), China (1), Denmark (4), Germany (10), England (3), Estonia (2), France (1), Greece (1), Japan (1), Luxembourg (1), Netherlands (1), New Zealand (2), Northern Ireland (1), Norway (1), Occupied Palestine (1), Portugal (1), Sweden (3), Switzerland (1), South Africa (1), and USA (1) (Albrecht, 2015).

[17] Sprint speed was measured in cross-sectional studies using 10 × 5 m pendulum running test in 424,328 Greek children and adolescents (49% girls; 6-18 years) in 2014 (Tambalis et al., 2016), 40 m sprint test among 6,398 children from Sweden, Germany, Hungary, Italy, Cyprus, Spain, Belgium and Estonia (3,288 girls; 6-10.9 years) from 2007 to 2008 (Miguel-Etayo et al., 2014), 50-m sprint and 4 x 10-m shuttle run in 3,804 Portuguese children (1,819 girls; 6-10 years) (Oliveira et al., 2014), 20 m sprint in 108,295 German children (8-9 years) (Fühner et al., 2021), and a meta-analytic approach with 15 studies using 50-m sprint in Australian children (5-18 years) from 1985 to 2009 (Catley & Tomkinson, 2013). Longitudinal studies assessed cardiorespiratory fitness in 240 German children (88 girls) aged 9-12 years at four consecutive annual assessments using 50-m sprint (Golle et al., 2014), and furthermore, analysed in a review of longitudinal studies from Australia (3), Belgium (4), Canada (3), China (1), Denmark (4), Germany (10), England (3), Estonia (2), France (1), Greece (1), Japan (1), Luxembourg (1), Netherlands (1), New Zealand (2), Northern Ireland (1), Norway (1), Occupied Palestine (1), Portugal (1), Sweden (3), Switzerland (1), South Africa (1), and USA (1) (Albrecht, 2015).

[18] Flexibility was assessed in cross-sectional studies using modified Back Saver Sit and Reach Test in 22,048 Portuguese children and adolescents (11,373 girls; 10-18 years) in 2008 (R. Santos et al., 2014), Sit and Reach Test in 424,328 Greek children and adolescents (49% girls; 6-18 years) in 2014 (Tambalis et al., 2016), sit and reach test in 3,804 Portuguese children (1,819 girls; 6-10 years) (Oliveira et al., 2014), Sit and Reach test in 3,742 German children and adolescents (50.1% girls; 4-23 years) from 2009 to 2013 (Niessner et al., 2020), Back Saver Sit and Reach Test on 6,398 children from Sweden, Germany, Hungary, Italy, Cyprus, Spain, Belgium and Estonia (3,288 girls; 6-10.9 years) from 2007 to 2008 (Miguel-Etayo et al., 2014), and another meta-analytic approach

with 15 studies using sit and reach test in Australian children (5-18 years) from 1985 to 2009 (Catley & Tomkinson, 2013). Longitudinal studies assessed cardiorespiratory fitness in 240 German children (88 girls) aged 9-12 years in four consecutive annual assessments using Stand and Reach Test (Golle et al., 2014), and also in a review of longitudinal studies from Australia (3), Belgium (4), Canada (3), China (1), Denmark (4), Germany (10), England (3), Estonia (2), France (1), Greece (1), Japan (1), Luxembourg (1), Netherlands (1), New Zealand (2), Northern Ireland (1), Norway (1), Occupied Palestine (1), Portugal (1), Sweden (3), Switzerland (1), South Africa (1), and USA (1) (Albrecht, 2015).

[19] Balance was assessed in cross-sectional studies using postural control with a one leg balance test and a blindfolded one leg balance test in 436 Swedish children (190 girls; 6-12 years) from 1999 to 2001 (Lundgren et al., 2011), static stance and backward balancing in 3,742 German children and adolescents (50.1% girls; 4-23 years) from 2009 to 2013 (Niessner et al., 2020), and a one-leg balance test in 6,398 children from Sweden, Germany, Hungary, Italy, Cyprus, Spain, Belgium and Estonia (3,288 girls; 6-10.9 years) from 2007 to 2008 (Miguel-Etayo et al., 2014).

[20] Several studies have used **global** as a description for trends (Tomkinson, 2007; Tomkinson et al., 2020; Tomkinson & Olds, 2007). However, as analysed data sets in these studies are heavily skewed towards European, North American, and Australasian countries, **international** was chosen to account for few/missing trends in South American, African, and Asian countries, as well as countries and areas outside these classifications (e.g., Peninsular, Arctic, and Antarctic).

21

countries	fitness tests	number of children	age range [years]	time period	citation
Australia, Belgium, Canada, France, Greece, Italy, the Netherlands, Northern Ireland, Spain, USA, and Poland	20 m shuttle run	129,882 (63,079 girls)	6 to 19	1980 to 2000	(Tomkinson et al., 2003)
Australia, Belgium, Bulgaria, Canada, China, Czech Republic, Estonia, Finland, France, Germany, Japan, Hong Kong, Hungary, occupied Palestine, Italy, Korean Republic, Lithuania, Mozambique, Netherlands, New Zealand, Northern Ireland, Poland, Russia, Singapore, Spain, Sweden, and USA, representing the geographical regions Africa/Middle-East (n = 2,542), Asia (n = 23,741,524), Australasia (n = 144,110), Europe (n = 995,950), and North America (n = 304,068)	20 m shuttle run, 5 to 12 min runs, and 300 to 2,414 m runs	25,455,527	6 to 19	1958 to 2003	(Tomkinson & Olds, 2007)
Australia, Belgium, Brazil, Canada, Estonia, France, Greece, Hungary, Italy, Japan, Lithuania, the Netherlands, Poland, Portugal, Seychelles, South Africa, Spain, UK, and USA	20 m shuttle run	965,264	9 to 17	1981 to 2014	(Tomkinson et al., 2019)
Australia, Brazil, Canada, Czech Republic, Denmark, England, Finland, Belgium, Germany, Hungary, Lithuania, Netherlands, New Zealand, Norway, Portugal, Sweden, and USA	20 m shuttle run, 6 to 12 min runs, 1200 to 1600 m run, maximal cycle ergometer, PWC 170 cycle ergometer, and a submaximal cycle ergometer test	96,522	6 to 18	1972 to 2015	(Fühner et al., 2020)

22

countries	fitness tests	number of children	age range [years]	time period	citation
Australia, Brazil, Canada, Czech Republic, Denmark, England, Finland, Belgium, Germany, Hungary, Lithuania, Netherlands, New Zealand, Norway, Portugal, Sweden, and USA	leg lift, sit-up, push-up, bent arm hang, pull-up, arm-pull, bench-press, or a two-hand lift test	96,522	6 to 18	1972 to 2015	(Fühner et al., 2020)
Australia, Belgium, Bulgaria, Canada, China, Estonia, France, Greece, Hong Kong, Italy, Japan, Mexico, Mozambique, Poland, Spain, Thailand, Turkey, UK, and USA	hand grip strength	2,230,658	9 to 17	1967 to 2017	(Dooley et al., 2020)

23

countries	fitness tests	number of children	age range [years]	time period	citation
Australia, Belgium, Bulgaria, Canada, China, Czech Republic, Czechoslovaki, Estonia, Finland, France, Germany, Hungary, Iceland, occupied Palestine, Italy, Japan, Korea, Lithuania, Mozambique, New Zealand, Poland, Singapore, Spain, Sweden, Thailand, Turkey, and USA in five geographical regions Africa/Middle East (n = 2192), Asia (n = 18,080,023), Australasia (n = 115,528), Europe (n = 1,861,841), and North America (n = 39,937)	single jump tests such as standing broad jump, vertical jump, running broad jump, or running vertical jump	49,123,233	6 to19	1958 to 2003	(Tomkinson, 2007)
Australia, Belgium, Bulgaria, Canada, China, Czech Republic, Estonia, Finland, France, Germany, Greece, Iceland, occupied Palestine, Italy, Japan, Lithuania, Mozambique, New Zealand, Poland, Republic of Korea, Singapore, Slovakia, Slovenia, Spain, Taiwan, Thailand, Turkey, UK, and USA	standing long jump	10,940,801	9 to 17	1960 to 2013	(Tomkinson et al., 2020).

24

countries	fitness tests	number of children	age range [years]	time period	citation
Australia, Belgium, Brazil, Bulgaria, Canada, China, Estonia , Finland, France, Greece, Hong Kong, Iceland, occupied Palestine, Italy, Japan, Lithuania, Mozambique, New Zealand, Norway, Poland, Republic of Korea, Singapore, Slovakia, Slovenia, Spain, Sweden, Switzerland, Taiwan, Thailand, Turkey, and UK	sit up test	9,939,289	9 to 17	1964 to 2017	(Kaster et al., 2020)

25

countries	fitness tests	number of children	age range [years]	time period	citation
Australia, Belgium, Bulgaria, Canada, China, Czech Republic, Czechoslovaki, Estonia, Finland, France, Germany, Hungary, Iceland, occupied Palestine, Italy, Japan, Korea, Lithuania, Mozambique, New Zealand, Poland, Singapore, Spain, Sweden, Thailand, Turkey, and USA in five geographical regions	20 to 100 m sprint, 4 x 9 to 13.7 m, shuttle run 6 x 13.7 m shuttle run, or 10 x 5 m shuttle run	49,123,233	6 to 19	1958 to 2003	(Tomkinson, 2007)
Africa/Middle East (n = 8,855), Asia (n = 25,427,141), Australasia (n = 100,796), Europe (n = 2,164,526), North America (n = 216,925) Australia, Brazil, Canada, Czech Republic, Denmark, England, Finland, Belgium, Germany, Hungary, Lithuania, Netherlands, New Zealand, Norway, Portugal, Sweden, and USA	20 to 50 m sprints, and 4 × 9 m shuttle run, or 10 × 5 m shuttle run	96,522	6 to 18	1972 to 2015	(Fühner et al., 2020)

[26] Secular trends were calculated for geographical regions Africa/Middle East, Asia, Australasia, Europe, and North America (Tomkinson, 2007; Tomkinson & Olds, 2007).

[27] Shuttle run performance trends (Tomkinson et al., 2019) were correlated with data from the Global Burden of Disease Study 2013, which provided trends for prevalences of overweight/obese children and adolescents (according to published cut off values (Cole et al., 2000; Cole & Lobstein, 2012)) from 183 countries from 1980 to 2013 (Ng et al., 2014). Trends for handgrip strength (Dooley et al., 2020), standing long jump (Tomkinson et al., 2020), and sit up performance (Kaster et al., 2020) were correlated with data from the NCD Risk Factor Collaboration which provided trends for mean BMI change (per decade) for 31.5 million boys and girls aged 5–19 years in 200 countries from 1975 to 2016 (Bentham et al., 2017).

[28] Physical activity trends from the Health Behaviour in School-aged Children survey were included in the studies. Physical activity was operationalised as change in prevalence of daily moderate to vigorous physical

activity for at least 60 minutes and vigorous physical activity for at least 60 minutes four times a week for boys and girls aged 11, 13 and 15 years in European countries between 2002 and 2014 (Inchley et al., 2017).

[29] Gross national income per capita in this analysis represents annually added value created by production of goods and services in a country (OECD, 2023) and was taken from the World Bank data set (Bank, 2023b). Classifications into low, middle, or high income economies were done using the World Bank atlas method (Bank, 2023b)

[30] The Gini index provides information on distributions of income within the population of a country. It is calculated by comparing the cumulative welfare share of households with the cumulative population share, where values of 0 represent perfect income equality and values of 100 represent perfect income inequality (Bank, 2023a). Data was retrieved from the World Bank data set (Bank, 2023b).

[31] Urbanisation provides information on percentage growth of urban populations and is calculated using population estimates of the World Bank and urban quotients from the World Urbanization Prospects of the United Nations (Bank, 2023a). Data was retrieved from the World Bank data set (Bank, 2023b).

[32] The human development index is composed of (1) life expectancy at birth (to measure ability to live a long and healthy life), (2) average number of years of schooling and expected number of years of schooling (to measure ability to acquire knowledge), and (3) gross national income per capita (to measure ability to achieve an adequate standard of living) (UNDP, 2018). Data was retrieved from the World Bank data set (Bank, 2023b).

[33] The federal state of Brandenburg, Germany surrounds Berlin and has a higher population density in close proximity to Berlin. The reasoning behind the analyses was that a higher population density is associated with an environment that offers more opportunities to improve physical fitness. Furthermore, average cost of living is higher in areas closer to Berlin. Accordingly, it is likely that more families with higher incomes live in this area.

[34] Overweight and obesity in children and adolescents have been defined using gender- and age-specific BMI values (Cole et al., 2000; Cole & Lobstein, 2012).

[35] Development status was defined in accordance with the *Global Burden of Diseases, Injuries, and Risk Factors* enterprise, based on gross national product per capita (Murray et al., 2012). Which countries were classified as “developed” or “developed” in which time period was not specified in the study or supplementary materials (Ng et al., 2014).

[36] Using age-specific growth references (i.e., medians and standard deviations) for children and adolescents, obesity was defined at the 2 standard deviation threshold (Onis et al., 2007).

[37] Body fat percent was calculated from skinfold thickness (assessed at triceps and subscapula) using Slaughter equations (Slaughter et al., 1988). Industrialised countries were Argentina, Australia, Bahrain, Belgium, Canada, the Czech Republic, Estonia, Finland, France, Germany, Greece, Hong Kong, Ireland, Italy, Japan, the Netherlands, New Zealand, Poland, Portugal, Russia, South Africa, Spain, Slovakia, South Africa, Sweden, Switzerland, Taiwan, Turkey, the United Arab Emirates, UK, and USA (Olds, 2009)

[38] Overweight and obese was assessed as (1) BMI (adjusted and age-specific), (2) skinfold thickness (triceps, subscapula, and sum of both), and (3) percentage body fat (estimated using skinfold thickness and Slaughter equations) in 6725 and 6866 children aged 5 to 17 years in the Bogalusa Heart Study from 1981 to 1994 (Freedman et al., 2009, 2013).

[39] Depressive symptoms were evaluated using depression rating scales such as the Child Depression Inventory, the Strength and Difficulties Questionnaire, and the Montgomery-Asberg Depression Rating Scale. Multiple definitions for overweight and obesity, such as BMI at 85th to 95th percentile for overweight and above the 95th percentile for obese, or age- & gender-specific curves, were included and pooled in the meta analysis (Quek et al., 2017).

[40] Obesity was defined as a BMI above the 75th percentile of the First National Health an Nutrition Examination Survey from 1971 to 1974 (USDA, 1990) for two consecutive years during their enrollment at Harvard School of Education (Must et al., 1992).

[41] Obesity was defined using BMI, which was included as a continuous variable in the analysis, or compared with age- and sex-specific BMI cut-off scores (Cole et al., 2000; Cole & Lobstein, 2012) to determine prevalences. Included studies examined subjects from Western Europe (Denmark, Finland, Norway, Sweden, UK, Netherlands), the USA (including Hawaii), Australia, and occupied Palestine (Park et al., 2012).

[42] Specific information on age clustering of secular BMI trends was not provided in this publication (Dooley et al., 2020).

[43] The study included 278 normal-weight children aged 5 to 15 years, where normal weight was defined as 2 SD above age-specific thresholds for mean BMI. The source for mean BMI was not reported. Percent body fat was estimated using a tetrapolar foot-to-foot impedance device.

[44] The study included 1224 boys and girls (i.e., ~768 normalweight, ~223 overweight, & ~233 obese) aged 6 to 15 and assessed weight status using the Centers for Disease Control and Prevention gender- and age-specific BMI reference values (Ogden & Flegal, 2010).

[45] The study included 1186 boys (992 normal weight, 194 overweight) and 1142 girls (1105 normal weight, 137 overweight) in 2003, and 661 boys (524 normal weight, 137 overweight) and 640 girls (563 normal weight, 77 overweight) in 2010. Weight status was defined using age- and gender-specific BMI cut-off points (Cole et al., 2000), and physical activity was assessed using a questionnaire, with adolescents grouped into inactive, mildly active, moderately active, active, and very active (Nupponen et al., 2010).

[46] The study analysed 270 boys and 249 girls at ages 7 to 15 years and defined weight status (i.e., thin, normalweight, overweight, & obese) using the World Health Organization's BMI references for children and adolescents (Onis et al., 2007).

[47] Inclusion criteria for socioeconomic position were parental education, income and/or occupation, and weight status was defined as (1) reference to presence of overweight (obesity and/or overweight) or (2) some anthropometric parameter. The richest countries were identified based on (1) membership of the Organisation for Economic Co-operation and Development and (2) a gross national income per capita of more than $25,000 according to the International Monetary Fund for 2010. These included Australia, Austria, Belgium, Canada, Czech Republic, Denmark, Finland, France, Germany, Greece, Holland, Iceland, Ireland, Occupied Palestine, Italy, Japan, Korea, Luxembourg, New Zealand, Norway, Portugal, Slovenia, Spain, Sweden, Switzerland, United Kingdom, and the USA.

[48] The study included 2389 adolescents aged 10 to 16 years, 52% girls and 48% boys, or 77% White and 23% African American. Prevalence of overweight was determined using age- and sex-specific BMI reference values (Najjar & Rowland, 1987).

[49] Intersectionality describes intertwining of social systems of oppression such as race, or gender with bodies that are placed at intersections of one or more margins are subject to greater oppression, or left out of social counter-movements (Crenshaw, 1989).

[50] 1491 adolescents (49.6% boys and 49.4% Black) were enrolled in the study. Weight status was defined as normal weight (i.e., BMI < 85%), at risk for overweight (i.e., BMI between 85% and 95%) and overweight (i.e., BMI ≥ 95%) using age- and sex-specific reference values (Kuczmarski et al., 2002). Socioeconomic status was defined using parental education and combined pre-tax total household income for the last 12 months, and perceptions of social stratification were assessed using the MacArthur Subjective Social Status Scale (Goodman et al., 2001).

[51] Prevalence of overweight was determined by percentage BMI at the 95th percentile using fitted linear growth curve modelling, and low-income households were defined as < $25,000/year.

[52] Executive function describes cognitive processes underlying selection, planning, coordination, and monitoring of complex, goal-directed processes in areas of perception, memory, and action (Donnelly et al., 2016).

[53] Hillman et al. (2005) studied progressive shuttle run performance (Welk et al., 2002) in 600 children and recruited 24 children (9.6 years) from the top and bottom 10%. Attention and working memory resources were assessed using the visual odd-ball paradigm, which required children to quickly identify an infrequent stimulus while not responding to a frequent stimulus (Hillman et al., 2005). Hillmann et al. (2009) determined progressive shuttle run performances (Welk et al., 2002) in 38 children (aged 8 to 11 years) and divided them into equally sized groups (19 high-fit, 19 low-fit). Reaction accuracy was determined with the Eriksen-Flanker task (Eriksen & Eriksen, 1974), in which children had to react as quickly as possible to different series of letters (Hillman et al., 2009).

[54] The study included 263 children (145 boys, 118 girls). Cognitive flexibility was assessed with the trail making test (Reitan, 1971) and problem solving ability with the Tower of London (Shallice, 1982).

[55] Executive function was measured by calculating a composite score from the Stroop-Gold colour-word test [selective attention, response inhibition, self-control and mental speed], the trail making test [measures attention, speed of psychomotor execution and mental flexibility], verbal semantic fluency [initiation, efficient organisation of verbal recall and memory and self-monitoring], and the backward digit span [working memory function].

[56] Skog et al. (2020) used the term 'executive function' differently in that publication than in this thesis. They refer to performance in a memory task, whereas in this paragraph executive function is specified above, which includes all cognitive tests used by Skog et al. (2020). They included 54 adolescents (aged 17 ± 2 years; 35 girls; 19 boys), and executive functions were assessed using the Two-Back Task [working memory task], the One-Back Task [speed in short-term memory], the One-Card Learning Test [accuracy in visual memory], the Continuous Paired-Associate Learning Task [visual learning and memory], the Identification Task [speed in an attentional task], the Detection Test [psychomotor function and reaction time], and the Groton Maze Learning Test ['executive function'].

[57] The study involved 423 Dutch adolescents (aged 13.45 ± 0.43 years; 46.8% boys). Visual-spatial working memory was examined with forward and backward conditions of the computerised grid task (Nutley et al., 2009), verbal working memory with forward and backward conditions of the digit span task (Wechsler, 1991), and interference control with an adapted version of the attention network test (Fan et al., 2002).

[58] The study included 527 students aged 12 to 15 years (13.5 ± 1.0 years; 340 males; 187 females), including 29 borderline students or students with intellectual-relational disabilities. Working memory updating and inhibition were assessed with the random number generation test.

[59] The study included 263 children (145 boys, 118 girls). Cognitive flexibility was assessed with the trail making test (Reitan, 1971) and problem solving ability with the Tower of London (Shallice, 1982).

[60] The study involved 423 Dutch adolescents (aged 13.45 ± 0.43 years; 46.8% boys). Visual-spatial working memory was examined with forward and backward conditions of the computerised grid task (Nutley et al., 2009), verbal working memory with forward and backward conditions of the digit span task (Wechsler, 1991), and interference control with an adapted version of the attention network test (Fan et al., 2002).

[61] The study involved 423 Dutch adolescents (aged 13.45 ± 0.43 years; 46.8% boys). Visual-spatial working memory was examined with forward and backward conditions of the computerised grid task (Nutley et al., 2009), verbal working memory with forward and backward conditions of the digit span task (Wechsler, 1991), and interference control with an adapted version of the attention network test (Fan et al., 2002).

[62] The study included 263 children (145 boys, 118 girls). Cognitive flexibility was assessed with the trail making test (Reitan, 1971) and problem solving ability with the Tower of London (Shallice, 1982).

[63] The study included 527 students aged 12 to 15 years (13.5 ± 1.0 years; 340 males; 187 females), including 29 borderline students or students with intellectual-relational disabilities. Working memory updating and inhibition were assessed with the random number generation test.

[64] As most countries in the Global South have some form of compulsory school attendance, most children living in these areas can be reached. However, as school attendance is linked to legal status, children without occupational status as well as children without a permanent residence or with attested 'educational and school inability' are systematically excluded from this approach (see Gomolla (2015) for a detailed analysis of institutional discrimination in the German education system).

[65] Intervention details were published elsewhere in Spanish (Ardoy et al., 2010).

[66] Physical education lessons were taught by regular physical education teachers and content was not controlled to reflect 'real-life' conditions. Other features of the intervention were additional health education focusing on physical activity and healthy eating, 3-4 days of teacher training, and improvement of indoor facilities for physical education and outdoor playgrounds. Sixty-nine children were recruited from 18 public preschools (10 intervention children [222 boys and 186 girls] and 8 control children [142 boys and 144 girls]).

[67] Three physical education lessons for intervention and control groups were taught by class teachers, while two additional physical education lessons in the intervention group were taught by physical education teachers. Details of additional specific physical education lessons were not reported. In addition to two extra physical education lessons per week, the intervention group introduced short daily physical activity breaks and physical activity homework. Mean age of the first grade children was 6.9 ± 0.3 years and that of fifth grade children was 11.1 ± 0.5 years. The intervention group consisted of 297 children (131 first graders and 166 fifth graders) and the control group consisted of 205 children (91 first graders and 114 fifth graders). The study was conducted with Swiss children.

[68] Additional physical education was provided by classroom teachers and included general activities within the regular curriculum such as ball games, running, jumping, and climbing. The intervention group consisted of 49 girls and 80 boys, and the control group consisted of 50 girls and 53 boys. Knee extensor strength was determined using peak isokinetic torque at 60° and 180° for knee flexors and extensors.

[69] Daily physical education classes at experimental schools and regular physical education classes at two control schools were taught by certified physical education teachers, focusing on fundamental skills.

[70] Ardoy et al. (2011) assessed BMI, fat mass index, waist circumference, waist circumference to height ratio, fat-free mass, and fat-free mass index. Bugge et al. (2012) measured waist circumference. Kriemler et al. (2010) recorded BMI and waist circumference. Löfgren et al. (2013) included BMI, fat-free mass (total, legs, and arms), and fat mass (total, legs, and arms) in their analysis. Reed et al. (2013) included mean BMI percentile in their analysis.

[71] Bugge et al. (2012) assessed systolic blood pressure, insulin resistance, total cholesterol to high-density lipoprotein ratio, triglycerides, and a composite score for cardiovascular disease, and Kriemler et al. (2010) included systolic blood pressure, diastolic blood pressure, triglycerides, total cholesterol to high-density lipoprotein ratio, glucose, and a cardiovascular risk score in their analyses.

[72] Kriemer et al. (2010) found trends favorable to the intervention for skinfold thickness and BMI, and Reed et al. (2013) found similar trends for BMI percentiles in girls only. Löfgren et al. (2013) found inverse trends (i.e., in favor of controls) for fat mass (total, legs, and arms) in boys and girls and for body fat percentage and BMI in girls only. Of note, they also found an increase in lean body mass (total, legs, and arms) and body mass in girls, which could indicate a difference in maturation and thus greater growth and maturation development in girls in the intervention condition.

[73] Supplemental physical education was conducted according to the CHAMPS study protocol (Wedderkopp et al., 2012). Overweight and obesity were classified using sex-specific body fat percentage standards (Williams et al., 1992).

[74] Puzzle-solving abilities were assessed using the standardized progressive matrices test (Raven et al., 1998), and perceptual speed for recognizing different or identical line patterns was measured using the perceptual speed test (Salthouse, 1996).

[75] The intervention lasted 6 months and included 70 children (37 intervention, 33 control). Overweight/lean children were classified using age-specific cutoffs for BMI (Cole et al., 2000). Inhibition and updating of working memory were assessed using various indices of the random number generation task (Towse & McLachlan, 1999).

[76] The intervention lasted 12 weeks, and high-intensity interval training was performed on a twice-per-week schedule. Memory was assessed with the RIAS test, which required participants to memorize and recall 15 playing cards (Santamaría-Fernández & Fernández-Pinto, 2013), and selective attention and concentration were assessed with the d2 test, which required participants to identify as many specific letters as possible in a given amount of time (Seisdedos, 2012).

[77] Both high-intensity interval trainings were performed in the first 8 to 10 minutes of three weekly physical education sessions over 8 weeks. The aerobic high-intensity interval training focused on gross motor, cardiorespiratory exercises, while the resistance and aerobic high-intensity interval training further included bodyweight training exercises.

[78] See https://mbjs.brandenburg.de/corona-aktuell/chronologie.html

[79] Keyage children (age 8.6 ± .3 years; 51.1% girls) were classified at age 6 to 6.99 at enrollment on September 30. September, younger and older children (age 7.9 ± .2 years, 57.8% girls & age 9.3 ± .3 years, 41.6% girls; respectively) were classified as one year younger and older than keyage children, respectively.

[80] The study included 765 6th and 765 8th grade pre-pandemic children aged 11.3 ± .5 and 13.3 ± .5 years, and 853 6th and 853 8th grade pandemic children aged 11.4 ± .5 and 13.4 ± .5 years. Physical fitness was assessed using SLOfit measures consisting of a hand tapping test, standing long jump, polygon course backward, sit-ups, stand and reach, bent-arm hang, 60 m sprint, and 600 m run, in addition to providing an overall index of physical fitness (Jurak et al., 2020).

[81] The review evaluated papers published between January 1st, 2020 and May 1st, 2020. Accordingly, potential long-term effects and newer variants of the SARS-CoV-2 virus were not considered.

[82] See https://www.uni-potsdam.de/fileadmin/projects/emotikon/SMaRTER-Studie/SMaRTER-Studie_Trainingsintervention_2019.pdf

[83] See https://www.uni-potsdam.de/fileadmin/projects/emotikon/SMaRTER-Studie/SMaRTER-Studie_Trainingsintervention_2019.pdf

[84] The names of the homework exercises are originally in German. Accordingly, some language-specific meaning may be lost in translation into English.

[85] A list of schools employing specially trained remedial physical education teachers was provided by the MBJS.

[86] See this https://www.uni-potsdam.de/de/emotikon/index or Section 1.1.3.2 for more details on the EMOTIKON study.

[87] Once at the start of the SMaRTER study and once at the first assessment during the Covid 19 pandemic (i.e., t6).

[88] VO2max was determined using a spiroergometer and a standardised protocol with increasing speed and incline (Dubowy et al., 2008).

[89] In the shuttle run, participants had to run back and forth along a 20 m track, the pace being increased every minute by an acoustic signal until the participants could no longer maintain the pace.

[90] See http://www.tatool-web.com/#!/doc/lib-exp-simon.html

[91] Household income was estimated for each child. If a child lived in more than one household, income was averaged across all associated households.

[92] Average income was estimated by dividing the household income by all individuals contributing to the household income.

[93] The Lorenz curve plots the cumulative income share (on the y-axis) against the cumulative population share (on the x-axis). If the income distribution in the population is completely equal, the Lorenz curve assumes a 45° line (Bank, 2023a).

[94] The Gini index calculates the area between a horizontal 45° line (which means perfect equality) and the plotted Lorenz curve (see footnote above). Accordingly, a Gini index of 0 represents perfect equality and a value of 100 represents perfect inequality (Willis, 2020).

[95] If the values are adjusted for the thirteenth salary, which is a voluntary bonus salary that the employer can pay at the end of the year and which can reach up to a full salary, the average monthly household income was €4,643.54 and the income per capita was €1,844.92.

[96] Interventions were pooled by defining t0 and t2 as baseline assessments and t1 and t3 as post tests.

[97] Kotzamanidis (Kotzamanidis, 2006) implemented a 10-week plyometric jumping program (60 to 100 jumps per session) into two weekly physical education lessons, with 15 children (age 11.1 ±.5 years) participating in the intervention and 15 children (age 10.9 ± .7 years) serving as a regular physical education control group. Faigenbaum et al. (2009) implemented 10 to 15 minutes of plyometric training consisting of jumps, leaps, sprints and throws with increasing intensity into one physical education lesson per week for 9 weeks. 70 children (aged 8 to 11 years) were enrolled in the study, with 40 children participating in the intervention and 34 serving as a control group for regular physical education classes.

[98] The resistance training program consisted of 16 minutes of circuit training with chest exercises, rowing, up and down walking, triceps extension, biceps curl, rope jumping, crunches and bridging with increasing work times and decreasing rests between exercises as the program progressed, followed by 24 minutes of racing games. The program was framed by 5 minute warm ups and cool downs.

[99] Analyses were conducted with. (1) 672 Greek schoolchildren (boys: n = 348, age 12.2 ± 0.7 years; girls: n = 324, age 12.2 ± 0.5 years) on 20 m shuttle run, 40 m sprint, vertical jump, and handgrip strength tests, with linear and quadratic trends for age also included in the models (Nevill et al., 2009), (2) 1,995 girls and 1,669 boys aged 11 to 17 from three areas in central Peru on 12 min run, standing long jump, grip strength, curl-up, and 10 x 5 m pendulum run tests, with maturity offset also included in the models (Valdivia et al., 2015), (3) 2,385 girls and 2,175 boys aged 9-15 from Peru and Brazil in 12 min run, standing long jump, grip strength, and 10 x 5 m pendulum run tests, with a maturity offset and physical activity also included in the models (Silva et al., 2016), (4) 3,058 boys and 4,044 girls aged 11-14 years living in an urban or rural area in northern Italy in sit-up and standing long jump (Lovecchio et al., 2019), and (5) 556 European sedentary children aged 11-13 years (282 boys; 274 girls) using the Cooper test (Giuriato et al., 2020).

[100] Gross motor coordination was assessed in 245 children (122 girls) aged 6 to 9 years from the Azores Islands using the 'Körperkoordinationstest für Kinder', which consisted of a balance task, successive lateral jumps, hopping on one leg over an obstacle and moving platforms (Schilling & Kiphard, 1974).

[101] Metabolic syndrome is comprised of abdominal obesity, insulin resistance, disordered blood lipids, hypertension and glucose intolerance as symptoms that increase the risk of cardiovascular disease and diabetes.

Further, in the cutoff scores are calculated with an emphasis on sensitivity, meaning correctly identifying children with metabolic syndrome rather than without.

[102] Included aerobic assessment are the one-mile run, PACER Test, and the Walk test.

[103] Muscular strength can be assessed using the curl-up, trunk extension, and/or 90° push up test. Flexibility can be assessed using the back saver sit and reach and the shoulder stretch.

[104] Anthropometrical information can be measured using skinfold assessments, bioelectrical impedance, or the BMI

[105] Fitness assessments were done focusing on aerobic capacity, body composition, abdominal strength and endurance, trunk extensor strength and endurance, upper body strength and endurance, and flexibility.

[106] California standardized test in math and English language arts.

[107] 1325 fourth to seventh and 1410 sixth to ninth grade children.

[108] Fitness assessments were done focusing on aerobic capacity, body composition, abdominal strength and endurance, trunk extensor strength and endurance, upper body strength and endurance, and flexibility.

[109] California standardized test in math and English language arts. 1101325 fourth to seventh and 1410 sixth to ninth grade children.

[111] 460 boys aged 13 ± 1 years

[112] Socioeconomic status was assesses using the eligibility for federal free lunch program with a low socioeconimic status being assigned to children eligible for the program, while all other non eligible children were assigned to the high socioeconomic status group.

[113] 492 girls aged 13 ± 1 years

[114] Home neighbourhood socioeconomic status was defined as the percentage of households below the federal poverty threshold according to United States Census 2010 boundaries.

[115] Cumulative poverty risk score was calculated for every child growing up with an income-to-needs ratio below 1 and adjusted using the per capita index for poverty in the united states. The ratio recorded every six months from the children's birth until the most recent assessment.

References

ABC, H. (2005). *Digit symbol substitution test - operations manual*. https://healthabc.nia.nih.gov/sites/default/files/dsst_0.pdf

Abdelkarim, O., Ammar, A., Trabelsi, K., Cthourou, H., Jekauc, D., Irandoust, K., Taheri, M., Bös, K., Woll, A., Bragazzi, N. L., & Hoekelmann, A. (2020). Prevalence of underweight and overweight and its association with physical fitness in egyptian schoolchildren. *International Journal of Environmental Research and Public Health*, *17*. https://doi.org/10.3390/ijerph17010075

Abdullah, A., Wolfe, R., Stoelwinder, J. U., Courten, M. de, Stevenson, C., Walls, H. L., & Peeters, A. (2011). The number of years lived with obesity and the risk of all-cause and cause-specific mortality. *International Journal of Epidemiology*, *40*. https://doi.org/10.1093/ije/dyr018

Acheson, R. M. (1966). Maturation of the skeleton. In F. Falkner (Ed.), *Human development* (pp. 465–502). Saunders.

Ahmed, S. (2013). Strange encounters: Embodied others in post-coloniality. In *Strange Encounters: Embodied Others in Post-Coloniality*. https://doi.org/10.4324/9780203349700

Ainsworth, C. (2015). Sex redefined. In *Nature* (Vol. 518). https://doi.org/10.1038/518288a

Akram, D. S., Astrup, A. V., Atinmo, T., Boissin, J. L., Bray, G. A., Carroll, K. K., Chitson, P., Chun-ming, C., Dietz, W. H., Hill, J. O., Jéquier, E., Komodiki, C., Matsuzawa, Y., Mollentze, W. F., Moosa, K., Noor, M. I., Reddy, K. S., Seidell, J., Tanphaichitr, V., ... Zimmet, P. (2000). Obesity: Preventing and managing the global epidemic. In *World Health Organization - Technical Report Series*.

Albrecht, C. A. (2015). *Entwicklung und einflussfaktoren der entwicklung der motorischen leistungsfähigkeit im kindes- und jugendalter - befunde der MoMo-längsschnittstudie*. Pädagogischen Hochschule Karlsruhe.

Alcoff, L. M. (2006). Visible identities: Race, gender, and the self. In *Visible Identities: Race, Gender, and the Self*. https://doi.org/10.1093/0195137345.001.0001

Alday, P. (2022). *Palday/MixedModelsExtras.jl: v0.1.5* [Computer software]. https://doi.org/10.5281/zenodo.7139991

Allaire, J. J., Teague, C., Scheidegger, C., Xie, Y., & Dervieux, C. (2022). *Quarto (version 1.2)* [Computer software]. https://doi.org/10.5281/zenodo.5960048

Amt für Statistik. (2021). *Primäreinkommen und verfügbares einkommen der privaten haushalte im land brandenburg 1991 bis 2020*. Amt für Statistik.

Andersen, K. L., Seliger, V., Rutenfranz, J., & Skrobak-Kaczynski, J. (1976). Physical performance capacity of children in norway. *European Journal of Applied Physiology and Occupational Physiology*, *35*. https://doi.org/10.1007/bf00444656

Ardoy, D. N., Fernández-Rodríguez, J. M., Chillón, P., Artero, E. G., España-Romero, V., Jiménez-Pavón, D., Ruiz, J. R., Guirado-Escámez, C., Castillo, M. J., & Ortega, F. B. (2010). Physical fitness enhancement through education, EDUFIT study: Background, design, methodology and dropout analysis. *Revista Espanola de Salud Publica*, *84*. https://doi.org/10.1590/s1135-57272010000200004

Ardoy, D. N., Fernández-Rodríguez, J. M., Ruiz, J. R., Chillón, P., España-Romero, V., Castillo, M. J., & Ortega, F. B. (2011). Improving physical fitness in adolescents through a school-based intervention: The EDUFIT study. *Revista Española de Cardiología (English Edition)*, *64*. https://doi.org/10.1016/j.rec.2011.02.010

Armstrong, N., & Welsman, J. (2007). Aerobic fitness: What are we measuring? In *Medicine and Sport Science* (Vol. 50). https://doi.org/10.1159/000101073

Artero, E. G., Espaa-Romero, V., Castro-Piero, J., Ortega, F. B., Suni, J., Castillo-Garzon, M. J., & Ruiz, J. R. (2011). Reliability of field-based fitness tests in youth. In *International Journal of Sports Medicine* (Vol. 32). https://doi.org/10.1055/s-0030-1268488

Asmussen, E., & Heebøll-Nielsen, K. (1955). A dimensional analysis of physical performance and growth in boys. *Journal of Applied Physiology*, *7*. https://doi.org/10.1152/jappl.1955.7.6.593

Auguie, B. (2017). *Miscellaneous functions for "grid" graphics type, version 2.3* [Computer software].

Bank, W. (2023a). *Poverty and inequality platform methodology handbook*. https://datanalytics.worldbank.org/PIP-Methodology/

Bank, W. (2023b). *World bank data*. https://datacatalog.worldbank.org/search?q=&sort=last_updated_date%20desc

Barriuso, L., Miqueleiz, E., Albaladejo, R., Villanueva, R., Santos, J. M., & Regidor, E. (2015). Socioeconomic position and childhoodadolescent weight status in rich countries: A systematic review, 1990-2013. *BMC Pediatrics*, *15*. https://doi.org/10.1186/s12887-015-0443-3

Bastian, C. C. von, Souza, A. S., & Gade, M. (2016). No evidence for bilingual cognitive advantages: A test of four hypotheses. *Journal of Experimental Psychology: General*, *145*, 246–258. https://doi.org/10.1037/xge0000120

Bates, D. M., Kliegl, R., Vasishth, S., & Baayen, H. (2018). Parsimonious mixed models douglas bates. *ArXiv e-Print, Under Revision*.

Bates, D., Alday, P., Kleinschmidt, D., Calderón, J. B. S., Zhan, L., Bouchet-Valat, M., Noack, A., Arslan, A., Kelman, T., Baldassari, A., Ehinger, B., Karrasch, D., Saba, E., Quinn, J., Hatherly, M., Piibeleht, M., Mogensen, P. K., Babayan, S., & Yakir, L. G. (2023). *JuliaStats/MixedModels.jl: v4.8.2* [Computer software]. https://doi.org/10.5281/zenodo.7529836

Bates, D., Mächler, M., Bolker, B. M., & Walker, S. C. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*. https://doi.org/10.18637/jss.v067.i01

Benjamin, H. J., & Glow, K. M. (2003). Strength training for children and adolescents: What can physicians recommend? In *Physician and Sportsmedicine* (Vol. 31). https://doi.org/10.1080/00913847. 2003.11439938

Bentham, J., Cesare, M. D., Bilano, V., Bixby, H., Zhou, B., Stevens, G. A., Riley, L. M., Taddei, C., Hajifathalian, K., Lu, Y., Savin, S., Cowan, M. J., Paciorek, C. J., Chirita-Emandi, A., Hayes, A. J., Katz, J., Kelishadi, R., Kengne, A. P., Khang, Y. H., ... Cisneros, J. Z. (2017). Worldwide trends in body-mass index, underweight, overweight, and obesity from 1975 to 2016: A pooled analysis of 2416 population-based measurement studies in 128·9 million children, adolescents, and adults. *The Lancet*, *390*. https://doi.org/10.1016/S0140-6736(17)32129-3

Beunen, G. P., Rogol, A. D., & Malina, R. M. (2006). Indicators of biological maturation and secular changes in biological maturation. In *Food and Nutrition Bulletin* (Vol. 27). https://doi.org/10.1177/15648265060274s508

Beunen, G., & Malina, R. M. (1988). Growth and physical performance relative to the timing of the adolescent spurt. *Exercise and Sport Sciences Reviews*, *16*. https://doi.org/10.1249/00003677-198800160-00018

Beurskens, R., Muehlbauer, T., & Granacher, U. (2015). Association of dual-task walking performance and leg muscle quality in healthy children. *BMC Pediatrics*, *15*. https://doi.org/10.1186/s12887-015-0317-8

Bezanson, J., Edelman, A., Karpinski, S., & Shah, V. B. (2017). Julia: A fresh approach to numerical computing. In *SIAM Review* (Vol. 59) [Computer software]. https://doi.org/10.1137/141000671

Blackless, M., Charuvastra, A., Derryck, A., Fausto-Sterling, A., Lauzanne, K., & Lee, E. (2000). How sexually dimorphic are we? Review and synthesis. *American Journal of Human Biology*, *12*. https://doi.org/10.1002/(sici)1520-6300(200003/04)12:2%3C151::aid-ajhb1%3E3.3.co;2-6

Blair, S. N., Falls, H. B., & Pate, R. R. (1983). A new physical fitness test. *Physician and Sportsmedicine*, *11*. https://doi.org/10.1080/00913847.1983.11708510

Bohr, A. D., Brown, D. D., Laurson, K. R., Smith, P. J. K., & Bass, R. W. (2013). Relationship between socioeconomic status and physical fitness in junior high school students. *Journal of School Health*, *83*. https://doi.org/10.1111/josh.12063

Borck, R., & Gohl, N. (2021). Gentrification and affordable housing policies. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.3980336

Bormann, A. (2016). *Wissenschaftliche analyse im rahmen der implementierung des einbeinstandtests in der primarstufe im land brandenburg (EMOTIKON-studie)*. Universität Potsdam.

Bös, K. (2000). AST 6-11 allgemeiner sportmotorischer test für kinder von 6 bis 11 jahren. *Haltung Und Bewegung*, *20*, 5–16.

Bös, K., Albrecht, C., Mewes, N., Tittlbach, S., & Woll, A. (2017). Aktivität und fitness von kindern und jugendlichen: Unverzichtbar für die gesundheit: Ergebnisse der MoMo-studie. In R. Zimmer & I. Hunger (Eds.), *Gut Starten: Bewegung— Entwicklung—Diversität*. Schorndorf, Germany.

Bös, K., Worth, A., Opper, E., Oberger, J., Romahn, N., Wagner, M., Jekauc, D., Mess, F., & Woll, A. (2009). *Motorik-modul: Eine studie zur motorischen leistungsfähigkeit und körperlich-sportlichen aktivität von kindern und jugendlichen in deutschland*. Bundesministerium für Familie, Senioren, Frauen und Jugend.

Bowe, B., Xie, Y., & Al-Aly, Z. (2022). Acute and postacute sequelae associated with SARS-CoV-2 reinfection. *Nature Medicine*, *28*, 2398–2405. https://doi.org/https://doi.org/10.1038/s41591-022-02051-3

Box, G. E. P., & Cox, D. R. (1964). An analysis of transformations. *Journal of the Royal Statistical Society: Series B (Methodological)*, *26*. https://doi.org/10.1111/j.2517-6161.1964.tb00553.x

Brockman, R., Jago, R., Fox, K. R., Thompson, J. L., Cartwright, K., & Page, A. S. (2009). "Get off the sofa and go and play": Family and socioeconomic influences on the physical activity of 10 -11 year old children. *BMC Public Health*, *9*. https://doi.org/10.1186/1471-2458-9-253

Brodin, P. (2022). SARS-CoV-2 infections in children: Understanding diverse outcomes. In *Immunity* (Vol. 55). https://doi.org/10.1016/j.immuni.2022.01.014

Bubb, W. J. (1986). Relative leanness. In E. Howley (Ed.), *Health/Fitness Instructor's Handbook* (pp. 51–79). Human Kinetics Publishers, Inc.

Bugge, A., El-Naaman, B., Dencker, M., Froberg, K., Holme, I. M. K., McMurray, R. G., & Andersen, L. B. (2012). Effects of a three-year intervention: The copenhagen school child intervention study. *Medicine and Science in Sports and Exercise*, *44*. https://doi.org/10.1249/MSS.0b013e31824bd579

Bull, F. C., Al-Ansari, S. S., Biddle, S., Borodulin, K., Buman, M. P., Cardon, G., Carty, C., Chaput, J. P., Chastin, S., Chou, R., Dempsey, P. C., Dipietro, L., Ekelund, U., Firth, J., Friedenreich, C. M., Garcia, L., Gichu, M., Jago, R., Katzmarzyk, P. T., ... Willumsen, J. F. (2020). World health organization 2020 guidelines on physical activity and sedentary behaviour. In *British Journal of Sports Medicine* (Vol. 54). https://doi.org/10.1136/bjsports-2020-102955

Butler, J. (1990). *Gender trouble: Feminism and the subversion of identity*. Routledge, Chapman & Hall Inc.

Cameron, N. D. (2012). The human growth curve, canalization and catch-up growth. In *Human Growth and Development*. https://doi.org/10.1016/B978-0-12-383882-7.00001-5

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 Reports*, *100*, 126.

Cassilhas, R. C., Viana, V. A. R., Grassmann, V., Santos, R. T., Santos, R. F., Tufik, S., & Mello, M. T. (2007). The impact of resistance exercise on the cognitive function of the elderly. *Medicine and Science in Sports and Exercise*, *39*. https://doi.org/10.1249/mss.0b013e318060111f

Castro-Piñero, J., Artero, E. G., España-Romero, V., Ortega, F. B., Sjöström, M., Suni, J., & Ruiz, J. R. (2010). Criterion-related validity of field-based fitness tests in youth: A systematic review. In *British Journal of Sports Medicine* (Vol. 44, pp. 934–943). https://doi.org/10.1136/bjsm.2009.058321

Castro-Piñero, J., González-Montesinos, J. L., Mora, J., Keating, X. D., Girela-Rejón, M. J., Sjöström, M., & Ruiz, J. R. (2009). Percentile values for muscular strength field tests in children aged 6 to 17 years: Influence of weight status. *Journal of Strength and Conditioning Research*, *23*. https://doi.org/10.1519/JSC.0b013e3181b8d5c1

Castro-Piñero, J., Ortega, F. B., Keating, X. D., González-Montesinos, J. L., Sjöstrom, M., & Ruiz, J. R. (2011). Percentile values for aerobic performance running/walking field tests in children aged 6 to 17 years: Influence of weight status. *Nutricion Hospitalaria : Organo Oficial de La Sociedad Espanola de Nutricion Parenteral y Enteral*, *26*.

Catley, M. J., & Tomkinson, G. R. (2013). Normative health-related fitness values for children: Analysis of 85347 test results on 9-17-year-old australians since 1985. *British Journal of Sports Medicine*, *47*. https://doi.org/10.1136/bjsports-2011-090218

Chaddock, L., Erickson, K. I., Prakash, R. S., Voss, M. W., VanPatter, M., Pontifex, M. B., Hillman, C. H., & Kramer, A. F. (2012). A functional MRI investigation of the association between childhood aerobic fitness and neurocognitive control. *Biological Psychology*, *89*. https://doi.org/10.1016/j.biopsycho.2011.10.017

Chaddock, L., Pontifex, M. B., Hillman, C. H., & Kramer, A. F. (2011). A review of the relation of aerobic fitness and physical activity to brain structure and function in children. *Journal of the International Neuropsychological Society*. https://doi.org/10.1017/S1355617711000567

Chaddock-Heyman, L., Erickson, K. I., Holtrop, J. L., Voss, M. W., Pontifex, M. B., Raine, L. B., Hillman, C. H., & Kramer, A. F. (2014). Aerobic fitness is associated with greater white matter integrity in children. *Frontiers in Human Neuroscience*, *8*. https://doi.org/10.3389/fnhum.2014.00584

Chan, J. W. C., Eysenck, H. J., & Lynn, R. (1991). Reaction times and intelligence among hong kong children. *Perceptual and Motor Skills*, *72*. https://doi.org/10.2466/pms.1991.72.2.427

Chu, C. H., Chen, F. T., Pontifex, M. B., Sun, Y., & Chang, Y. K. (2019). Health-related physical fitness, academic achievement, and neuroelectric measures in children and adolescents. *International Journal of Sport and Exercise Psychology*, *17*. https://doi.org/10.1080/1612197X.2016.1223420

Chumela, W. C., Roche, A. F., & Thissen, D. (1989). The FELS method of assessing the skeletal maturity of the hand   wrist. *American Journal of Human Biology*, *1*. https://doi.org/10.1002/ajhb.1310010206

Chung, S. (2015). Body mass index and body composition scaling to height in children and adolescent. *Annals of Pediatric Endocrinology & Metabolism*, *20*. https://doi.org/10.6065/apem.2015.20. 3.125

Clare, E. (2015). *Exile & pride - disability, queerness, and liberation* (3rd ed.). Duke University Press.

Clemes, S. A., & Biddle, S. J. H. (2013). The use of pedometers for monitoring physical activity in children and adolescents: Measurement considerations. *Journal of Physical Activity and Health*, *10*. https://doi.org/10.1123/jpah.10.2.249

Cole, T. J., Bellizzi, M. C., Flegal, K. M., & Dietz, W. H. (2000). Establishing a standard definition for child overweight and obesity worldwide: International survey. *British Medical Journal*, *320*. https://doi.org/10.1136/bmj.320.7244.1240

Cole, T. J., & Lobstein, T. (2012). Extended international (IOTF) body mass index cut-offs for thinness, overweight and obesity. *Pediatric Obesity*, *7*. https://doi.org/10.1111/j.2047-6310.2012.00064.x

Collins, P. H. (2002). *Black feminist thought - knowledge, consciousness and the politics of empowerment* (2nd ed.). Routledge.

Collins, P. H., & Bilge, S. (2016). *Intersectionality*. Polity Press.
Corbin, C. B., Dowell, L. J., Lindsay, R., & Tolson, H. (1978). *Concepts in physical education with laboratories and experiments* (3rd ed.). W. C. Brown & Company, Publishers.

Corbin, C. B., Pangrazi, R. P., & Franks, D. B. (2000). Definitions : Health , fitness , and physical activity. *President's Council on Physical Fitness and Sports Research Digest*.

Costigan, S. A., Eather, N., Plotnikoff, R. C., Hillman, C. H., & Lubans, D. R. (2016). High-intensity interval training for cognitive and mental health in adolescents. *Medicine and Science in Sports and Exercise*, *48*. https://doi.org/10.1249/MSS.0000000000000993

Cotman, C. W., Berchtold, N. C., & Christie, L. A. (2007). Exercise builds brain health: Key roles of growth factor cascades and inflammation. In *Trends in Neurosciences* (Vol. 30). https://doi.org/10.1016/j.tins.2007.06.011

Crenshaw, K. (1989). Demarginalizing the intersection of race and sex: A black feminist critique of antidiscrimination doctrine, feminist theory and antiracist politics. University of chicago legal forum, 139, 139–167. *University of Chicago Legal Forum*, *78*.

Crova, C., Struzzolino, I., Marchetti, R., Masci, I., Vannozzi, G., Forte, R., & Pesce, C. (2014). Cognitively challenging physical activity benefits executive function in overweight children. *Journal of Sports Sciences*, *32*. https://doi.org/10.1080/02640414.2013.828849

Cureton, K. J., Plowman, S. A., & Mahar, M. T. (2013). FITNESSGRAM /ACTIVITYGRAM reference guide (4th edition). In *The Cooper Institute, Dallas, TX*.

Dalen, D. B. van, Mitchell, E. D., & Bennett, R. L. (1953). *A world history of physical education.* Prentice-Hall lnc.

Davidson, M. C., Amso, D., Anderson, L. C., & Diamond, A. (2006). Development of cognitive control and executive functions from 4 to 13 years: Evidence from manipulations of memory, inhibition, and task switching. *Neuropsychologia*, *44*, 2037–2078. https://doi.org/10.1016/j.neuropsychologia.2006.02.006

Deforche, B., Lefevre, J., Bourdeaudhuij, I. D., Hills, A. P., Duquet, W., & Bouckaert, J. (2003). Physical fitness and physical activity in obese and nonobese flemish youth. *Obesity Research*, *11*. https://doi.org/10.1038/oby.2003.59

Deprez, D., Coutts, A. J., Fransen, J., Deconinck, F., Lenoir, M., Vaeyens, R., & Philippaerts, R. (2013). Relative age, biological maturation and anaerobic characteristics in elite youth soccer players. *International Journal of Sports Medicine*, *34*. https://doi.org/10.1055/s-0032-1333262

Diamond, A. (2000). Close interrelation of motor development and cognitive development and of the cerebellum and prefrontal cortex. *Child Development*, *71*. https://doi.org/10.1111/1467-8624.00117

Diamond, A. (2006). The early development of executive functions. In E. Bialystock & F. Craik (Eds.), *Lifespan cognition* (pp. 70–95). Oxford University Press.

Diamond, A. (2013). Executive functions. *Annual Review of Psychology*, *64*, 135–168. https://doi.org/10.1146/annurev-psych-113011-143750

Diedrichsen, J., Criscimagna-Hemminger, S. E., & Shadmehr, R. (2007). Dissociating timing and coordination as functions of the cerebellum. *Journal of Neuroscience*, *27*. https://doi.org/10.1523/JNEUROSCI.0061-07.2007

DIW. (2017). *Wählerschaft der parteien*. Deutsches Institut für Wirtschaftsforschung. Donnelly, J. E., Hillman, C. H., Castelli, D., Etnier, J. L., Lee, S., Tomporowski, P., Lambourne, K., & Szabo-Reed, A. N. (2016). Physical activity, fitness, cognitive function, and academic achievement in children: A systematic review. In *Medicine and Science in Sports and Exercise* (Vol. 48). https://doi.org/10.1249/MSS.0000000000000901

Dooley, F. L., Kaster, T., Fitzgerald, J. S., Walch, T. J., Annandale, M., Ferrar, K., Lang, J. J., Smith, J. J., & Tomkinson, G. R. (2020). A systematic analysis of temporal trends in the handgrip strength of 2,216,320 children and adolescents between 1967 and 2017. In *Sports Medicine* (Vol. 50). https://doi.org/10.1007/s40279-020-01265-0

Drewnowski, A., & Eichelsdoerfer, P. (2010). The economics of obesity: Why are poor people fat? In *Obesity Prevention: The Role of Brain and Society on Individual Behavior*. https://doi.org/10.1016/B978-0-12-374387-9.00058-1

Dubowy, K. O., Baden, W., Bernitzki, S., & Peters, B. (2008). A practical and transferable new protocol for treadmill testing of children and adults. *Cardiology in the Young*, *18*. https://doi.org/10.1017/S1047951108003181

Dumith, S. C., Ramires, V. V., Souza, M. A., Moraes, D. S., Petry, F. G., Oliveira, E. S., Ramires, S. V., & Hallal, P. C. (2010). Overweight/obesity and physical fitness among children and adolescents. *Journal of Physical Activity and Health*, *7*. https://doi.org/10.1123/jpah.7.5.641

Elgar, F. J., Stefaniak, A., & Wohl, M. J. A. (2020). The trouble with trust: Time-series analysis of social capital, income inequality, and COVID-19 deaths in 84 countries. *Social Science and Medicine*, *263*. https://doi.org/10.1016/j.socscimed.2020.113365

Ellert, U., Ravens-Sieberer, U., Erhart, M., & Kurth, B. M. (2011). Determinants of agreement between self-reported and parent-assessed quality of life for children in germany-results of the german health interview and examination survey for children and adolescents (KiGGS). *Health and Quality of Life Outcomes*, *9*. https://doi.org/10.1186/1477-7525-9-102

Elliot, A. J. (2019). A historically based review of empirical work on color and psychological functioning: Content, methods, and recommendations for future research. *Review of General Psychology*, *23*. https://doi.org/10.1037/gpr0000170
Elliot, A. J., & Aarts, H. (2011). Perception of the color red enhances the force and velocity of motor output. *Emotion*, *11*. https://doi.org/10.1037/a0022599
Erhart, M., Ellert, U., Kurth, B. M., & Ravens-Sieberer, U. (2009). Measuring adolescents' HRQoL via self reports and parent proxy reports: An evaluation of the psychometric properties of both versions of the KINDL-r instrument. *Health and Quality of Life Outcomes*, *7*. https://doi.org/10.1186/1477-7525-7-77
Eriksen, B. A., & Eriksen, C. W. (1974). Effects of noise letters upon the identification of a target letter in a nonsearch task. *Perception & Psychophysics*, *16*. https://doi.org/10.3758/BF03203267
Ervin, R. B., Fryar, C. D., Wang, C. Y., Miller, I. M., & Ogden, C. L. (2014). Strength and body weight in US children and adolescents. *Pediatrics*, *134*. https://doi.org/10.1542/peds.2014-0794
Esmaeilzadeh, S. (2014). Relationship between depressive symptoms with physical activity and physical fitness among children. *Mental Health and Prevention*, *2*. https://doi.org/10.1016/j.mhp.2014.05.002
Esmaeilzadeh, S. (2015). The association between depressive symptoms and physical status including physical activity, aerobic and muscular fitness tests in children. *Environmental Health and Preventive Medicine*, *20*. https://doi.org/10.1007/s12199-015-0484-0
Ettman, C. K., Abdalla, S. M., Cohen, G. H., Sampson, L., Vivier, P. M., & Galea, S. (2021). Low assets and financial stressors associated with higher depression during COVID-19 in a nationally representative sample of US adults. *Journal of Epidemiology and Community Health*, *75*. https://doi.org/10.1136/jech-2020-215213
Evans, G. W. (2003). A multimethodological analysis of cumulative risk and allostatic load among rural children. *Developmental Psychology*, *39*. https://doi.org/10.1037/0012-1649.39.5.924
Faigenbaum, A., & Farrell, A. (2009). "Plyo play": A novel program of short bouts of moderate and high intensity exercise improves physical fitness in elementary school children the physiology and biomechanics of load carriage performance view project. *Physical Educator*, *66*, 37. https://www.researchgate.net/publication/282026495
Fan, J., McCandliss, B. D., Sommer, T., Raz, A., & Posner, M. I. (2002). Testing the efficiency and independence of attentional networks. *Journal of Cognitive Neuroscience*, *14*. https://doi.org/10. 1162/089892902317361886
Fancourt, D., Steptoe, A., & Bu, F. (2020). Trajectories of depression and anxiety during enforced isolation due to COVID-19: Longitudinal analyses of 59,318 adults in the UK with and without diagnosed mental illness. *medRxiv*.
Fanon, F. (1986). *Black skin, white mask*. Pluto Press.
Faude, O., Nowacki, P. E., & Urhausen, A. (2004). Comparison of non-invasive tests to assess cardiorespiratory fitness in school children. *DEUTSCHE ZEITSCHRIFT FUR SPORTMEDIZIN*, *55*.
Fausto-Sterling, A. (2020). *Sexing the body - gender politics and the construction of sexuality* (2nd ed.). Basic Books - Hachette Book Group.
Fausto-Sterling, A., & Trajanoski, Ž. (2004). The five sexes: Why male and female are not enough. *Identities: Journal for Politics, Gender and Culture*, *3*. https://doi.org/10.51151/identities.v3i1.118
Fernandez-Santos, J. R., Ruiz, J. R., Cohen, D. D., Gonzalez-Montesinos, J. L., & Castro-Piñero, J. (2015). Reliability and validity of tests to assess lower-body muscular power in children. *Journal of Strength and Conditioning Research*, *29*, 2277–2285. https://doi.org/10.1519/JSC.0000000000000864
Figura, F., Cama, G., Capranica, L., Guidetti, L., & Pulejo, C. (1991). Asessment of static balance in children. *Journal of Sports Medicine and Physical Fitness*, *31*.
Fjørtoft, I., Pedersen, A. V., Sigmundsson, H., & Vereijken, B. (2011). Measuring physical fitness in children who are 5 to 12 years old with a test battery that is functional and easy to administer. *Physical Therapy*, *91*, 1087–1095. https://doi.org/10.2522/ptj.20090350

Fox, J., & Weisberg, S. (2019). An {r} companion to applied regression, third edition. In *Thousand Oaks CA: Sage.* [Computer software]. https://CRAN.R-project.org/package=flextable

Fredricks, J. A., & Eccles, J. S. (2005). Family socialization, gender, and sport motivation and involvement. *Journal of Sport and Exercise Psychology*, *27*. https://doi.org/10.1123/jsep.27.1.3

Freedman, D. S., Horlick, M., & Berenson, G. S. (2013). A comparison of the slaughter skinfoldthickness equations and BMI in predicting body fatness and cardiovascular disease risk factor levels in children1-4. *American Journal of Clinical Nutrition*, *98*. https://doi.org/10.3945/ajcn.113.065961

Freedman, D. S., Katzmarzyk, P. T., Dietz, W. H., Srinivasan, S. R., & Berenson, G. S. (2009). Relation of body mass index and skinfold thicknesses to cardiovascular disease risk factors in children: The bogalusa heart study. *American Journal of Clinical Nutrition*, *90*. https://doi.org/10.3945/ajcn. 2009.27525

Fühner, T., Granacher, U., Golle, K., & Kliegl, R. (2021). Age and sex effects in physical fitness components of 108,295 third graders including 515 primary schools and 9 cohorts. *Scientific Reports*, *11*. https://doi.org/10.1038/s41598-021-97000-4

Fühner, T., Granacher, U., Golle, K., & Kliegl, R. (2022). Effect of timing of school enrollment on physical fitness in third graders. *Nature - Scientific Reports*, *12*.

Fühner, T., Kliegl, R., Arntz, F., Kriemler, S., & Granacher, U. (2020). An update on secular trends in physical fitness of children and adolescents from 1972 to 2015: A systematic review. In *Sports Medicine* (pp. 1–18). Springer Science; Business Media Deutschland GmbH. https://doi.org/10.1007/s40279-020-01373-x

Garciá-Hermoso, A., Ramírez-Vélez, R., Lubans, D. R., & Izquierdo, M. (2021). Effects of physical education interventions on cognition and academic performance outcomes in children and adolescents: A systematic review and meta-analysis. *British Journal of Sports Medicine*, *55*. https://doi.org/10.1136/bjsports-2021-104112

García-Hermoso, A., Alonso-Martínez, A. M., Ramírez-Vélez, R., Pérez-Sousa, M. Á., Ramírez-Campillo, R., & Izquierdo, M. (2020). Association of physical education with improvement of health-related physical fitness outcomes and fundamental motor skills among youths. *JAMA Pediatrics*, e200223. https://doi.org/10.1001/jamapediatrics.2020.0223

García-Hermoso, A., Ramírez-Campillo, R., & Izquierdo, M. (2019). Is muscular fitness associated with future health benefits in children and adolescents? A systematic review and meta-analysis of longitudinal studies. In *Sports Medicine* (Vol. 49). https://doi.org/10.1007/s40279-019-01098-6

Garland-Thomson, R. (2011). Misfits: A feminist materialist disability concept. *Hypatia*, *26*. https://doi.org/10.1111/j.1527-2001.2011.01206.x

Gerodimos, V. (2012). Reliability of handgrip strength test in basketball players. *Journal of Human Kinetics*, *31*. https://doi.org/10.2478/v10078-012-0003-y

Gilman, S. L. (1994). *Difference and pathology - stereotypes of sexuality, race and madness* (5th ed.). Cornell University.

Giuriato, M., Nevill, A., Kawczynski, A., & Lovecchio, N. (2020). A cross-sectional study of cooper test in children according to allometric scaling. *Journal of Sports Medicine and Physical Fitness*, *60*. https://doi.org/10.23736/S0022-4707.19.10282-4

Giuriato, M., Pugliese, L., Biino, V., Bertinato, L., Torre, A. L., & Lovecchio, N. (2019). Association between motor coordination, body mass index, and sports participation in children 6–11 years old. *Sport Sciences for Health*, *15*. https://doi.org/10.1007/s11332-019-00554-0

Gohel, D. (2022). *Officer: Manipulation of microsoft word and PowerPoint documents, version 0.4.4* [Computer software].

Gohel, D., & Skintzos, P. (2022). *Flextable: Functions for tabular reporting, version 0.8.3* [Computer software].

Golle, K. (2015). *Physical fitness in school-aged children*. University of Potsdam.

Golle, K., Granacher, U., Hoffmann, M., Wick, D., & Muehlbauer, T. (2014). Effect of living area and sports club participation on physical fitness in children: A 4 year longitudinal study. *BMC Public Health*, *14*. https://doi.org/10.1186/1471-2458-14-499

Golle, K., Muehlbauer, T., Wick, D., & Granacher, U. (2015). Physical fitness percentiles of german children aged 9-12 years: Findings from a longitudinal study. *PLoS ONE*, *10*. https://doi.org/10.1371/journal.pone.0142393

Gómez-Costilla, P., García-Prieto, C., & Somarriba-Arechavala, N. (2022). Aging and gender health gap: A multilevel analysis for 17 european countries. *Social Indicators Research*, *160*. https://doi.org/10.1007/s11205-020-02595-2

Gomolla, M. (2015). Institutionelle diskriminierung im bildungs- und erziehungssystem. In R. Leiprecht & A. Steinbach (Eds.), *Schule in der Migrationsgesellschaft. Ein Handbuch. Band 1: Grundlagen – Diversität – Fachdidaktiken* (pp. 193–219). Debus Pädagogik.

Goodman, E., Adler, N. E., Daniels, S. R., Morrison, J. A., Slap, G. B., & Dolan, L. M. (2003). Impact of objective and subjective social status on obesity in a biracial cohort of adolescents. *Obesity Research*, *11*. https://doi.org/10.1038/oby.2003.140

Goodman, E., Adler, N. E., Kawachi, I., Frazier, A. L., Huang, B., & Colditz, G. A. (2001). Adolescents' perceptions of social status: Development and evaluation of a new indicator. *Pediatrics*, *108*. https://doi.org/10.1542/peds.108.2.e31

Goyder, J., Warriner, K., & Miller, S. (2002). Evaluating socio-economic status (SES) bias in survey nonresponse. *Journal of Official Statistics*, *18*.

Granacher, U., & Golle, K. (2016). *Generierung von normwerten und die prüfung der reliabilität des einbeinstands- und standweitsprungtests*. Universität Potsdam.

Greulich, W. W., & Pyle, S. I. (1959). *Radiographic atlas of skeletal development of the hand and wrist* (2nd ed.). Stanford University Press.

Haapala, E. A. (2013). Cardiorespiratory fitness and motor skills in relation to cognition and academic performance in children - a review. In *Journal of Human Kinetics* (Vol. 36). https://doi.org/10.2478/hukin-2013-0006

Haaren, B. von, Härtel, S., Seidel, I., Schlenker, L., & Bös, K. (2011). Validity of a 6-min endurance run and a 20-m shuttle run in 9- to 11-year old children. *Deutsche Zeitschrift Fur Sportmedizin*, *62*.

Hall, P. A., Meng, G., Hudson, A., Sakib, M. N., Hitchman, S. C., MacKillop, J., Bickel, W. K., & Fong, G. T. (2022). Cognitive function following SARS-CoV-2 infection in a population-representative canadian sample. *Brain, Behavior, and Immunity - Health*, *21*. https://doi.org/10.1016/j.bbih.2022.100454

Haverkamp, B. F., Oosterlaan, J., Königs, M., & Hartman, E. (2021). Physical fitness, cognitive functioning and academic achievement in healthy adolescents. *Psychology of Sport and Exercise*, *57*. https://doi.org/10.1016/j.psychsport.2021.102060

Hawkins, D., & Metheny, J. (2001). Overuse injuries in youth sports: Biomechanical considerations. *Medicine and Science in Sports and Exercise*, *33*. https://doi.org/10.1097/00005768-200110000-00014

Haywood, K. M., & Getchell, N. (2014). *Life span motor development* (M. Schrag, C. M. Drews, S. Huls, & M. J. Zavala, Eds.; 6th ed.). Human Kinetics.

Heisig, J. P. (2021). Soziale ungleichheit und gesundheitliches risiko in der pandemie. In *Corona. Pandemie und Krise* (pp. 332–344). Bundeszentrale für politische Bildung.

Herdt, G. (1990). Mistaken gender: 5-alpha reductase hermaphroditism and biological reductionism in sexual identity reconsidered. *American Anthropologist*, *92*. https://doi.org/10.1525/aa.1990.92.2.02a00110

Herdt, G. H., & Davidson, J. (1988). The sambia "turnim-man": Sociocultural and clinical aspects of gender formation in male pseudohermaphrodites with 5-alpha-reductase deficiency in papua new guinea. *Archives of Sexual Behavior*, *17*. https://doi.org/10.1007/BF01542051

Hilchey, M. D., & Klein, R. M. (2011). Are there bilingual advantages on nonlinguistic interference tasks? Implications for the plasticity of executive control processes. In *Psychonomic Bulletin and Review* (Vol. 18). https://doi.org/10.3758/s13423-011-0116-7

Hillman, C. H., Buck, S. M., Themanson, J. R., Pontifex, M. B., & Castelli, D. M. (2009). Aerobic fitness and cognitive development: Event-related brain potential and task performance indices of executive control in preadolescent children. *Developmental Psychology*, *45*. https://doi.org/10.1037/a0014437

Hillman, C. H., Castelli, D. M., & Buck, S. M. (2005). Aerobic fitness and neurocognitive function in healthy preadolescent children. *Medicine and Science in Sports and Exercise*, *37*. https://doi.org/10.1249/01.mss.0000176680.79702.ce
Hillman, C. H., Erickson, K. I., & Kramer, A. F. (2008). Be smart, exercise your heart: Exercise effects on brain and cognition. In *Nature Reviews Neuroscience* (Vol. 9). https://doi.org/10.1038/nrn2298
Hochberg, Z., & Belsky, J. (2013). Evo-devo of human adolescence: Beyond disease models of early puberty. In *BMC Medicine* (Vol. 11). https://doi.org/10.1186/1741-7015-11-113
Hohenstein, S., & Kliegl, R. (2022). *Remef: Remove partial effects, version 1.0.7* [Computer software]. https://github.com/hohenstein/remef/
hooks, bell. (1990). Yearning race, gender and cultural politics. In *Routledge* (Vol. 1).
Hotelling, H. (1933). Analysis of a complex of statistical variables into principal components. *Journal of Educational Psychology*, *24*. https://doi.org/10.1037/h0071325
Howley, E. T., & Franks, D. B. (1986). *Health/fitness instructor's handbook*. Human Kinetics Publisher Inc.
Hunsicker, P. A., & Reiff, G. G. (1976). AAHPER youth fitness test manual. *American Alliance for'Health, Physical Education, and Recreation, Washington, D.C.*
Hyland-Monks, R., Cronin, L., McNaughton, L., & Marchant, D. (2018). The role of executive function in the self-regulation of endurance performance: A critical review. In *Progress in Brain Research* (Vol. 240). https://doi.org/10.1016/bs.pbr.2018.09.011
Inchley, J., Currie, D., Jewell, J., Breda, J., & Barnekow, V. (2017). Adolescent obesity and related behaviours: Trends and inequalities in the WHO region 2002-2014. *World Health Organization, Regional Office for Europe*.
Ishihara, T., Sugasawa, S., Matsuda, Y., & Mizuno, M. (2018). Relationship between sports experience and executive function in 6–12-year-old children: Independence from physical fitness and moderation by gender. *Developmental Science*, *21*. https://doi.org/10.1111/desc.12555
Izquierdo-Pujol, J., Moron-Lopez, S., Dalmau, J., Gonzalez-Aumatell, A., Carreras-Abad, C., Mendez, M., Rodrigo, C., & Martinez-Picado, J. (2022). Post COVID-19 condition in children and adolescents: An emerging problem. *Frontiers in Pediatrics*, *10*. https://doi.org/10.3389/fped.2022.894204
Jacobsen, G. M., Mello, C. M. de, Kochhann, R., & Fonseca, R. P. (2017). Executive functions in school-age children: Influence of age, gender, school type and parental education. *Applied Cognitive Psychology*, *31*. https://doi.org/10.1002/acp.3338
Jekauc, D., Reimers, A. K., Wagner, M. O., & Woll, A. (2013). Physical activity in sports clubs of children and adolescents in germany: Results from a nationwide representative survey. *Journal of Public Health (Germany)*, *21*. https://doi.org/10.1007/s10389-013-0579-2
Joensuu, L., Syväoja, H., Kallio, J., Kulmala, J., Kujala, U. M., & Tammelin, T. H. (2018). Objectively measured physical activity, body composition and physical fitness: Cross-sectional associations in 9- to 15-year-old children. *European Journal of Sport Science*, *18*. https://doi.org/10.1080/17461391.2018.1457081
Johnson, J. L., & Repta, R. (2012). Sex and gender: Beyond the binaries. In *Designing and Conducting Gender, Sex, and Health Research*. https://doi.org/10.4135/9781452230610.n2
Jordan-Young, R. M. (2010). *Brainstorm - the flaws in the science of sex differences*. Harvard University Press.
Jordan-Young, R., & Karkazis, K. (2012). Some of their parts " gender verification " and elite sports. *Anthropology News*, *53*, 1–34.
Jordan-Young, R., & Karkazis, K. (2019). *Testosterone: An unauthorized biography*. Harvard University Press.
Jurak, G., Cooper, A., Leskošek, B., & Kovač, M. (2013). Long-term effects of 4-year longitudinal school-based physical activity intervention on the physical fitness of children and youth during 7-year follow-up assessment. *Central European Journal of Public Health*, *21*. https://doi.org/10.21101/cejph.a3823
Jurak, G., Leskošek, B., Kovač, M., Sorić, M., Kramaršič, J., Sember, V., Đurić, S., Meh, K., Morrison, S. A., Strel, J., & Starc, G. (2020). SLOfit surveillance system of somatic and

motor development of children and adolescents: Upgrading the slovenian sports educational chart. *AUC KINANTHROPOLOGICA*, *56*. https://doi.org/10.14712/23366052.2020.4
Kanehisa, H., Yata, H., Ikegawa, S., & Fukunaga, T. (1995). A cross-sectional study of the size and strength of the lower leg muscles during growth. *European Journal of Applied Physiology and Occupational Physiology*, *72*. https://doi.org/10.1007/BF00964130
Karkazis, K., & Jordan-Young, R. M. (2018). The powers of testosterone: Obscuring race and regional bias in the regulation of women athletes. *Feminist Formations*, *30*. https://doi.org/10.1353/ff.2018. 0017
Karkazis, K., Jordan-Young, R., Davis, G., & Camporesi, S. (2012). Out of bounds? A critique of the new policies on hyperandrogenism in elite female athletes. *American Journal of Bioethics*, *12*. https://doi.org/10.1080/15265161.2012.680533
Kaster, T., Dooley, F. L., Fitzgerald, J. S., Walch, T. J., Annandale, M., Ferrar, K., Lang, J. J., Smith, J. J., & Tomkinson, G. R. (2020). Temporal trends in the sit-ups performance of 9,939,289 children and adolescents between 1964 and 2017. *Journal of Sports Sciences*. https://doi.org/10.1080/02640414. 2020.1763764
Kessler, S. J., & McKenna, W. (1978). *Gender: An ethnomethodological approach*. Wiley.
Kharel, M., Sakamoto, J. L., Carandang, R. R., Ulambayar, S., Shibanuma, A., Yarotskaya, E., Basargina, M., & Jimba, M. (2022). Impact of COVID-19 pandemic lockdown on movement behaviours of children and adolescents: A systematic review. In *BMJ Global Health* (Vol. 7). https://doi.org/10.1136/bmjgh-2021-007190
Klakk, H., Chinapaw, M., Heidemann, M., Andersen, L. B., & Wedderkopp, N. (2013). Effect of four additional physical education lessons on body composition in children aged 8-13 years - a prospective study during two school years. *BMC Pediatrics*, *13*. https://doi.org/10.1186/1471-2431-13-170
Kliegl, R. M. (1982). *On relations between cognition and reading style: Individual differences and developmental trends*. University of Colorado.
Kliegl, R., & Teich, P. (2022). *Analysen in EMOTIKON: Belege für sozialstrukturelle korrelationen mit schuleingangsuntersuchungen (SEU) und der bundestagswahl 2021*. Fachkonferenzleiter Sport der Schulen mit Grundschulteil in Potsdam-Mittelmark.
KMK. (1999). *Grundsätze für die durchführung von sportförderunterricht sowie für die ausbildung und prüfung zum erwerb der befähigung für das erteilen von sportförderunterricht*. Kultusministerkonferenz.
Konty, K. J., Day, S. E., Larkin, M., Thompson, H. R., & D'Agostino, E. M. (2020). Physical fitness disparities among new york city public school youth using standardized methods, 2006-2017. *PLoS ONE*, *15*. https://doi.org/10.1371/journal.pone.0227185
Kotzamanidis, C. (2006). Effect of plyometric training on running performance and vertical jumping in prepubertal boys. *Journal of Strength and Conditioning Research*, *20*, 441–445.
Kovacs, V. A., Starc, G., Brandes, M., Kaj, M., Blagus, R., Leskošek, B., Suesse, T., Dinya, E., Guinhouya, B. C., Zito, V., Rocha, P. M., Gonzalez, B. P., Kontsevaya, A., Brzezinski, M., Bidiugan, R., Kiraly, A., Csányi, T., & Okely, A. D. (2022). Physical activity, screen time and the COVID-19 school closures in europe–an observational study in 10 countries. *European Journal of Sport Science*, *22*. https://doi.org/10.1080/17461391.2021.1897166
Koziol, L. F., Budding, D., Andreasen, N., D'Arrigo, S., Bulgheroni, S., Imamizu, H., Ito, M., Manto, M., Marvel, C., Parker, K., Pezzulo, G., Ramnani, N., Riva, D., Schmahmann, J., Vandervert, L., & Yamazaki, T. (2014). Consensus paper: The cerebellum's role in movement and cognition. In *Cerebellum* (Vol. 13). https://doi.org/10.1007/s12311-013-0511-x
Kraus, H., & Hirschland, R. P. (1954). Minimum muscular fitness tests in school children. *Research Quarterly of the American Association for Health, Physical Education and Recreation*, *25*. https://doi.org/10.1080/10671188.1954.10624957
Kriemler, S., Meyer, U., Martin, E., Sluijs, E. M. F. V., Andersen, L. B., & Martin, B. W. (2011). Effect of school-based interventions on physical activity and fitness in children and adolescents: A review of reviews and systematic update. In *British Journal of Sports Medicine* (Vol. 45, pp. 923–930). https://doi.org/10.1136/bjsports-2011-090186
Kriemler, S., Zahner, L., Schindler, C., Meyer, U., Hartmann, T., Hebestreit, H., Rocca, H. P. B.-L., Mechelen, W. van, & Puder, J. J. (2010). Effect of school based physical activity programme

(KISS) on fitness and adiposity in primary schoolchildren: Cluster randomised controlled trial. *BMJ (Clinical Research Ed.)*. https://doi.org/10.1136/bmj.c785
Kromeyer-Hauschild, K., Wabitsch, M., Kunze, D., Geller, F., Geiß, H. C., Hesse, V., Hippel, A. von, Jaeger, U., Johnsen, D., Korte, W., Menner, K., Müller, G., Müller, J. M., Niemann-Pilatus, A., Remer, T., Schaefer, F., Wittchen, H.-U., Zabransky, S., Zellner, K., ... Hebebrand, J. (2001). Perzentile für den body-mass-index für das kindes- und jugend-alter unter heranziehung verschiedener deutscher stichproben. *Monatsschrift Kinderheilkunde*, *149*, 807–818.
Kuczmarski, R. J., Ogden, C. L., Guo, S. S., Grummer-Strawn, L. M., Flegal, K. M., Mei, Z., Wei, R., Curtin, L. R., Roche, A. F., & Johnson, C. L. (2002). 2000 CDC growth charts for the united states: Methods and development. *Vital and Health Statistics, Series 11: Data from the National Health Survey, the National Health and Nutrition Examination Surveys, and the Hispanic Health and Nutrition Examination Survey*, *11*.
Kunin-Batson, A. S., Crain, L., Gunnar, M. R., Kelly, A. S., Kharbanda, E. O., Haapala, J., Seburg, E. M., Sherwood, N. E., & French, S. A. (2023). Household income, cortisol, and obesity during early childhood: A prospective longitudinal study. *The Journal of Pediatrics*, *252*, 76–82.
Kvalø, S. E., Dyrstad, S. M., Bru, E., & BrØnnick, K. (2019). Relationship between aerobic fitness and academic performance: The mediational role of executive function. *Journal of Sports Medicine and Physical Fitness*, *59*. https://doi.org/10.23736/S0022-4707.18.08971-5
Lai, S. K., Costigan, S. A., Morgan, P. J., Lubans, D. R., Stodden, D. F., Salmon, J., & Barnett, L. M. (2014). Do school-based interventions focusing on physical activity, fitness, or fundamental movement skill competency produce a sustained impact in these outcomes in children and adolescents? A systematic review of follow-up studies. In *Sports Medicine* (Vol. 44, pp. 67–79). https://doi.org/10.1007/s40279-013-0099-9
Lampert, T., Müters, S., Stolzenberg, H., & Kroll, L. E. (2014). Messung des sozioökonomischen status in der KiGGS-studie: Erste folgebefragung (KiGGS welle 1). *Bundesgesundheitsblatt - Gesundheitsforschung - Gesundheitsschutz*, *57*. https://doi.org/10.1007/s00103-014-1974-8
Lauretis, T. de. (1987). Technologies of gender: Essays on theory, film, and fiction. In *Gender Theories in the Modern World: An Interdisciplinary Approach* (pp. 1–30).
Lawrenz, W., & Stemper, T. (2012). Comparison of 6-minute-jog-walk and maximal oxygen uptake in 8-10-year old school children. *DEUTSCHE ZEITSCHRIFT FUR SPORTMEDIZIN*, *63*.
Lechuga, J. R., Molina, J. J. M., Sánchez, J. M., Muñoz, C. S., Marzo, P. F., & Díaz, M. Z. (2012). Effect of an 8-week aerobic training program during physical education lessons on aerobic fitness in adolescents. *Nutricion Hospitalaria*, *27*. https://doi.org/10.3305/nh.2012.27.3.5725
Lee, P. A., Nordenström, A., Houk, C. P., Ahmed, S. F., Auchus, R., Baratz, A., Dalke, K. B., Liao, L. M., Lin-Su, K., Looijenga, L. H. J., Mazur, T., Meyer-Bahlburg, H. F. L., Mouriquand, P., Quigley, C. A., Sandberg, D. E., Vilain, E., & Witchel, S. (2016). Global disorders of sex development update since 2006: Perceptions, approach and care. In *Hormone research in paediatrics* (Vol. 85). https://doi.org/10.1159/000442975
Léger, L. A., Mercier, D., Gadoury, C., & Lambert, J. (1988). The multistage 20 metre shuttle run test for aerobic fitness. *Journal of Sports Sciences*, *6*. https://doi.org/10.1080/02640418808729800
Li, J., Bünning, M., Kaiser, T., & Hipp, L. (2022). Who suffered most? Parental stress and mental health during the COVID-19 pandemic in germany. *Journal of Family Research*, *34*. https://doi.org/10.20377/jfr-704
Li, S. C. (2003). Biocultural orchestration of developmental plasticity across levels: The interplay of biology and culture in shaping the mind and behavior across the life span. In *Psychological Bulletin* (Vol. 129). https://doi.org/10.1037/0033-2909.129.2.171
Li, S. C., Lindenberger, U., Hommel, B., Aschersleben, G., Prinz, W., & Baltes, P. B. (2004). Transformations in the couplings among intellectual abilities and constituent cognitive processes across the life span. *Psychological Science*. https://doi.org/10.1111/j.0956-7976.2004.01503003.x
Liguoro, I., Pilotto, C., Bonanni, M., Ferrari, M. E., Pusiol, A., Nocerino, A., Vidal, E., & Cogo, P. (2020). SARS-COV-2 infection in children and newborns: A systematic review. In *European Journal of Pediatrics* (Vol. 179). https://doi.org/10.1007/s00431-020-03684-7

Lindsay, R. S., Hanson, R. L., Roumain, J., Ravussin, E., Knowler, W. C., & Tataranni, P. A. (2001). Body mass index as a measure of adiposity in children and adolescents: Relationship to adiposity by dual energy x-ray absorptiometry and to cardiovascular risk factors. In *Journal of Clinical Endocrinology and Metabolism* (Vol. 86). https://doi.org/10.1210/jcem.86.9.7760

Lobstein, T., Baur, L., & Uauy, R. (2004). Obesity in children and young people: A crisis in public health. In *Obesity Reviews, Supplement* (Vol. 5). https://doi.org/10.1111/j.1467-789x.2004.00133.x

Löfgren, B., Daly, R. M., Nilsson, J. Å., Dencker, M., & Karlsson, M. K. (2013). An increase in school-based physical education increases muscle strength in children. In *Medicine and Science in Sports and Exercise* (Vol. 45). https://doi.org/10.1249/MSS.0b013e31827c0889

London, R. A., & Castrechini, S. (2011). A longitudinal examination of the link between youth physical fitness and academic achievement. *Journal of School Health*, *81*. https://doi.org/10.1111/j.1746-1561.2011.00608.x

Lopes, V. P., Malina, R. M., Maia, J. A. R., & Rodrigues, L. P. (2018). Body mass index and motor coordination: Non-linear relationships in children 6–10 years. *Child: Care, Health and Development*, *44*. https://doi.org/10.1111/cch.12557

Lovecchio, N., Giuriato, M., Zago, M., & Nevill, A. (2019). Identifying the optimal body shape and composition associated with strength outcomes in children and adolescent according to place of residence: An allometric approach. *Journal of Sports Sciences*, *37*. https://doi.org/10.1080/02640414.2018.1562615

Lovell, R., Towlson, C., Parkin, G., Portas, M., Vaeyens, R., & Cobley, S. (2015). Soccer player characteristics in english lower-league development programmes: The relationships between relative age, maturation, anthropometry and physical fitness. *PLoS ONE*, *10*. https://doi.org/10.1371/journal.pone.0137238

Lüdecke, D. (2021). *sjPlot: Data visualization for statistics in social science, r package version version 2.8.12*.

Lüdecke, D., Ben-Shachar, M., Patil, I., Waggoner, P., & Makowski, D. (2021). Performance: An r package for assessment, comparison and testing of statistical models. *Journal of Open Source Software*, *6*. https://doi.org/10.21105/joss.03139

Lugones, M. (2010). Toward a decolonial feminism. *Hypatia*, *25*. https://doi.org/10.1111/j.1527-2001.2010.01137.x

Lundgren, S. S., Nilsson, J. Å., Ringsberg, K. A. M., & Karlsson, M. K. (2011). Normative data for tests of neuromuscular performance and DXA-derived lean body mass and fat mass in pre-pubertal children. *Acta Paediatrica, International Journal of Paediatrics*, *100*. https://doi.org/10.1111/j. 1651-2227.2011.02322.x

Ma, X. (2000). Socioeconomic gaps in academic achievement within schools: Are they consistent across subject areas? *International Journal of Phytoremediation*, *21*. https://doi.org/10.1076/edre. 6.4.337.6935

Malina, R. (1988). Human growth: Selected aspects of current research on well-nourished children. *Annual Review of Anthropology*, *17*. https://doi.org/10.1146/annurev.anthro.17.1.187

Malina, R. M. (2004). Secular trends in growth, maturation and physical performance : A review. *Anthropological Review*, *67*.

Malina, R. M., & Katzmarzyk, P. T. (2006). Physical activity and fitness in an international growth standard for preadolescent and adolescent children. In *Food and Nutrition Bulletin* (Vol. 27). https://doi.org/10.1177/15648265060274s511

Manna, I. (2014). Growth development and maturity in children and adolescent: Relation to sports and physical activity. *American Journal of Sports Science and Medicine*, *2*. https://doi.org/10.12691/ajssm-2-5a-11

Marchetti, R., Forte, R., Borzacchini, M., Vazou, S., Tomporowski, P. D., & Pesce, C. (2015). Physical and motor fitness, sport skills and executive function in adolescents: A moderated prediction model. *Psychology*, *06*. https://doi.org/10.4236/psych.2015.614189

Marmot, M., & Allen, J. (2020). COVID-19: Exposing and amplifying inequalities. In *Journal of Epidemiology and Community Health* (Vol. 74). https://doi.org/10.1136/jech-2020-214720

Marshall, W. A., & Tanner, J. M. (1969). Variations in pattern of pubertal changes in girls. *Archives of Disease in Childhood*, *44*. https://doi.org/10.1136/adc.44.235.291

Marshall, W. A., & Tanner, J. M. (1970). Variations in the pattern of pubertal changes in boys. *Archives of Disease in Childhood*, *45*. https://doi.org/10.1136/adc.45.239.13

Martínez-Lopez, E. J., Torre-Cruz, M. J. de la, Suárez-Manzan, S., & Ruiz-Ariza, A. (2018). 24 sessions of monitored cooperative high-intensity interval training improves attention-concentration and mathematical calculation in secondary school. *Journal of Physical Education and Sport*, *18*. https://doi.org/10.7752/jpes.2018.03232

Marx, K. (1955). *The poverty of philosophy*. Progress Publishers.

MBJS. (2011). *Handreichung sport - empfehlung zur umsetzung - sportförderunterricht*. Ministerium für Bildung, Jugend und Sport - Brandenburg.

MBJS. (2020a). *3. Bericht des ministeriums für bildung, jugend und sport zum aktuellen sachstand bei den getroffenen maßnahmen zur eindämmung des corona-virus zu TOP 1 der 6. Sitzung des ausschusses für bildung, jugend und sport des landtages brandenburg am 20. Mai 2020*. Ministerium für Bildung, Jugend und Sport - Brandenburg.

MBJS. (2020b). *4. Bericht des ministeriums für bildung, jugend und sport zum aktuellen sachstand bei den getroffenen maßnahmen zur eindämmung des corona-virus zu TOP 4 der 7. Sitzung des ausschusses für bildung, jugend und sport des landtages brandenburg am 4. Juni 2020*. Ministerium für Bildung, Jugend und Sport - Brandenburg.

MBJS. (2020c). *Bericht des ministeriums für bildung, jugend und sport zu den getroffenen maßnahmen im geschäftsbereich des ministeriums zur eindämmung der corona-pandemie – aktueller sachstand zu TOP 1 der 4. Sitzung des ausschusses für bildung, jugend und sport des landtages brandenburg am 26. März 2020*. Ministerium für Bildung, Jugend und Sport - Brandenburg.

MBJS. (2021a). *18. Sitzung des ABJS am 7. Mai 2021 TOP 2 bericht des ministeriums für bildung, jugend und sport zum aktuellen sachstand bei den getroffenen maßnahmen zur eindämmung des corona-virus, insbesondere in den bereichen schule und kindertagesstätten sowie zum aktuellen sachstand bei der umsetzung des digitalpakts schule*. Ministerium für Bildung, Jugend und Sport - Brandenburg.

MBJS. (2021b). *20. Sitzung des ausschusses für bildung, jugend und sport am 12. August 2021 bericht des ministeriums für bildung, jugend und sport zum aktuellen sachstand bei den getroffenen maßnahmen zur eindämmung des coronavirus, insbesondere in den bereichen schule und kindertagesstätte sowie zum aktuellen sachstand bei der umsetzung des digitalpakts schule*. Ministerium für Bildung, Jugend und Sport - Brandenburg.

MBJS. (2021c). *25. Sitzung des ausschusses für bildung, jugend und sport, 2. Dezember 2021 bericht des ministeriums für bildung, jugend und sport (MBJS) zum aktuellen sachstand bei den getroffenen maßnahmen zur eindämmung des coronavirus, insbesondere in den bereichen schule und kindertagesstätte sowie zum aktuellen sachstand bei der umsetzung des DigitalPakts schule*. Ministerium für Bildung, Jugend und Sport - Brandenburg.

MBJS. (2022a). *26. Sitzung des ausschusses für bildung, jugend und sport, 6. Januar 2022 bericht des ministeriums für bildung, jugend und sport (MBJS) zum aktuellen sachstand bei den getroffenen maßnahmen zur eindämmung des coronavirus, insbesondere in den bereichen schule und kindertagesbetreuung sowie zum aktuellen sachstand bei der umsetzung des DigitalPakts schule*. Ministerium für Bildung, Jugend und Sport - Brandenburg.

MBJS. (2022b). *28. Sitzung des ausschusses für bildung, jugend und sport, 10. März 2022 bericht des ministeriums für bildung, jugend und sport (MBJS) zum aktuellen sachstand bei den getroffenen maßnahmen zur eindämmung des coronavirus, insbesondere in den bereichen schule und kindertagesbetreuung sowie zum aktuellen sachstand bei der umsetzung des DigitalPakts schule und des aktionsprogramms „aufholen nach corona für kinder und jugendliche"*. Ministerium für Bildung, Jugend und Sport - Brandenburg.

MBJS. (2022c). *30. Sitzung des ausschusses für bildung, jugend und sport am 5. Mai 2022 bericht des ministeriums für bildung, jugend und sport zum aktuellen sachstand bei den getroffenen maßnahmen zur eindämmung des coronavirus, insbesondere in den bereichen schule und kindertagesstätte, sowie zum aktuellen sachstand bei der umsetzung des digitalpakts schule*. Ministerium für Bildung, Jugend und Sport - Brandenburg.

MBJS. (2022d). *31. Sitzung des ausschusses für bildung, jugend und sport am 9. Juni 2022 bericht des ministeriums für bildung, jugend und sport zum aktuellen sachstand bei den getroffenen* *maßnahmen zur eindämmung des coronavirus, insbesondere in den bereichen schule und kindertagesstätte, sowie zum aktuellen sachstand bei der umsetzung des digitalpakts schule*. Ministerium für Bildung, Jugend und Sport - Brandenburg.

McMurray, R. G., Harrell, J. S., Deng, S., Bradley, C. B., Cox, L. M., & Bangdiwala, S. I. (2000). The influence of physical activity, socioeconomic status, and ethnicity on the weight status of adolescents. *Obesity Research*, *8*. https://doi.org/10.1038/oby.2000.14

McRuer, R., & Berube, M. (2006). Crip theory: Cultural signs of queerness and disability. In *Crip Theory: Cultural Signs of Queerness and Disability*. https://doi.org/10.1080/15017410701880122

Mehta, N. S., Mytton, O. T., Mullins, E. W. S., Fowler, T. A., Falconer, C. L., Murphy, O. B., Langenberg, C., Jayatunga, W. J. P., Eddy, D. H., & Nguyen-Van-Tam, J. S. (2020). SARS-CoV-2 (COVID-19): What do we know about children? A systematic review. In *Clinical Infectious Diseases* (Vol. 71). https://doi.org/10.1093/cid/ciaa556

Mendívil, E. R., & Sarbo, B. (2022). *Die diversität der ausbeutung - zur kritik des herrschenden antirassismus*. Dietz.

Meyer-Bahlburg, H. F. L., Dalke, K. B., Berenbaum, S. A., Cohen-Kettenis, P. T., Hines, M., & Schober, J. M. (2016). Gender assignment, reassignment and outcome in disorders of sex development: Update of the 2005 consensus conference. *Hormone Research in Paediatrics*, *85*. https://doi.org/10.1159/000442386

Miguel-Etayo, P. D., Gracia-Marco, L., Ortega, F. B., Intemann, T., Foraita, R., Lissner, L., Oja, L., Barba, G., Michels, N., Tornaritis, M., Molnár, D., Pitsiladis, Y., Ahrens, W., & Moreno, L. A. (2014). Physical fitness reference standards in european children: The IDEFICS study. *International Journal of Obesity*, *38*. https://doi.org/10.1038/ijo.2014.136

Mirwald, R. L., Baxter-Jones, A. D. G., Bailey, D. A., & Beunen, G. P. (2002). An assessment of maturity from anthropometric measurements. *Medicine and Science in Sports and Exercise*, *34*, 689–694. https://doi.org/10.1249/00005768-200204000-00020

Mischler, F. (2021). *THE GENDER PAY GAP a cause analysis based on the structure of earnings survey 2018* (pp. 110–125). Federal Statistical Office (Statistisches Bundesamt).

Molloy, E. J., Nakra, N., Gale, C., Dimitriades, V. R., & Lakshminrusimha, S. (2022). Multisystem inflammatory syndrome in children (MIS-c) and neonates (MIS-n) associated with COVID-19: Optimizing definition and management. *Pediatric Research*, 1–10.

Moore, S. A., McKay, H. A., Macdonald, H., Nettlefold, L., Baxter-Jones, A. D. G., Cameron, N., & Brasher, P. M. A. (2015). Enhancing a somatic maturity prediction model. *Medicine and Science in Sports and Exercise*, *47*. https://doi.org/10.1249/MSS.0000000000000588

Murray, C. J. L., Ezzati, M., Flaxman, A. D., Lim, S., Lozano, R., Michaud, C., Naghavi, M., Salomon, J. A., Shibuya, K., Vos, T., Wikler, D., & Lopez, A. D. (2012). GBD 2010: Design, definitions, and metrics. In *The Lancet* (Vol. 380). https://doi.org/10.1016/S0140-6736(12)61899-6

Must, A., Jacques, P. F., Dallal, G. E., Bajema, C. J., & Dietz, W. H. (1992). Long-term morbidity and mortality of overweight adolescents. *New England Journal of Medicine*, *327*. https://doi.org/10.1056/nejm199211053271904

Nagata, J. M., Magid, H. S. A., & Gabriel, K. P. (2020). Screen time for children and adolescents during the coronavirus disease 2019 pandemic. *Obesity*, *28*. https://doi.org/10.1002/oby.22917

Najjar, M. F., & Rowland, M. (1987). Anthropometric reference data and prevalence of overweight, united states, 1976-80. *Vital and Health Statistics. Series 11, Data from the National Health Survey*.

Nevill, A., Tsiotra, G., Tsimeas, P., & Koutedakis, Y. (2009). Allometric associations between body size, shape, and physical performance of greek children. *Pediatric Exercise Science*, *21*. https://doi.org/10.1123/pes.21.2.220

Ng, M., Fleming, T., Robinson, M., Thomson, B., Graetz, N., Margono, C., Mullany, E. C., Biryukov, S., Abbafati, C., Abera, S. F., Abraham, J. P., Abu-Rmeileh, N. M. E., Achoki, T., Albuhairan, F. S., Alemu, Z. A., Alfonso, R., Ali, M. K., Ali, R., Guzman, N. A., ... Gakidou, E. (2014). Global, regional, and national prevalence of overweight and obesity in children and

adults during 1980-2013: A systematic analysis for the global burden of disease study 2013. *The Lancet*, *384*. https://doi.org/10.1016/S0140-6736(14)60460-8
Niessner, C., Granacher, U., & Woll, A. (2021). Entwicklung von kraft, ausdauer, schnelligkeit, beweglichkeit und koordination. In *Pädiatrische Sportmedizin*. https://doi.org/10.1007/978-3-662-61588-1_2
Niessner, C., Utesch, T., Oriwol, D., Hanssen-Doose, A., Schmidt, S. C. E., Woll, A., Bös, K., & Worth, A. (2020). Representative percentile curves of physical fitness from early childhood to early adulthood: The MoMo study. *Frontiers in Public Health*, *8*. https://doi.org/10.3389/fpubh.2020.00458
Niet, A. G. van der, Hartman, E., Smith, J., & Visscher, C. (2014). Modeling relationships between physical fitness, executive functioning, and academic achievement in primary school children. *Psychology of Sport and Exercise*, *15*. https://doi.org/10.1016/j.psychsport.2014.02.010
Nieuwenhuis, J., & Xu, J. (2021). Residential segregation and unequal access to schools. *Social Inclusion*, *9*. https://doi.org/10.17645/si.v9i2.3606
Nishida, C., Barba, C., Cavalli-Sforza, T., Cutter, J., Deurenberg, P., Darnton-Hill, I., Deurenberg-Yap, M., Gill, T., James, P., Ko, G., Kosulwat, V., Kumanyika, S., Kurpad, A., Mascie-Taylor, N., Moon, H. K., Nakadomo, F., Nishida, C., Noor, M. I., Reddy, K. S., ... Zimmet, P. (2004). Appropriate body-mass index for asian populations and its implications for policy and intervention strategies. *The Lancet*, *363*. https://doi.org/10.1016/S0140-6736(03)15268-3
Norman, A. C., Drinkard, B., McDuffie, J. R., Ghorbani, S., Yanoff, L. B., & Yanovski, J. A. (2005). Influence of excess adiposity on exercise fitness and performance in overweight children and adolescents. *Pediatrics*, *115*. https://doi.org/10.1542/peds.2004-1543
Nupponen, H., Laakso, L., Rimpelä, A., Pere, L., & Telama, R. (2010). Questionnaire-assessed moderate to vigorous physical activity of the finnish youth in 1979-2005. *Scandinavian Journal of Medicine and Science in Sports*, *20*. https://doi.org/10.1111/j.1600-0838.2009.00875.x
Nutley, S. B., Söderqvist, S., Bryde, S., Humphreys, K., & Klingberg, T. (2009). Measuring working memory capacity with greater precision in the lower capacity ranges. *Developmental Neuropsychology*, *35*. https://doi.org/10.1080/87565640903325741
Nuyens, F. M., Kuss, D. J., Lopez-Fernandez, O., & Griffiths, M. D. (2019). The empirical analysis of non-problematic video gaming and cognitive skills: A systematic review. In *International Journal of Mental Health and Addiction* (Vol. 17). https://doi.org/10.1007/s11469-018-9946-0
Nuzzo, J. L. (2020). The case for retiring flexibility as a major component of physical fitness. *Sports Medicine*, *50*. https://doi.org/10.1007/s40279-019-01248-w
Nygaard, U., Holm, M., Hartling, U. B., Glenthøj, J., Schmidt, L. S., Nordly, S. B., Matthesen, A. T., Linstow, M.-L. von, & Espenhain, L. (2022). Incidence and clinical phenotype of multisystem inflammatory syndrome in children after infection with the SARS-CoV-2 delta variant by vaccination status: A danish nationwide prospective cohort study. , 6(7), 459-465. *The Lancet Child & Adolescent Health*, *6*, 459–465.
Oberger, J., Romahn, N., Opper, E., Tittlbach, S., Wank, V., Woll, A., Worth, A., & Bös, K. (2006). Untersuchungen zur motorischen leistungsfähigkeit und körperlich-sportlichen aktivität im rahmen des kinder- und jugendgesundheitssurveys des robert koch-institutes berlin. In G. Wydra, H. Winchenbach, M. Schwarz, & K. Pfeifer (Eds.), *Assessmentverfahren in Gesundheitssport und Bewegungstherapie : Messen, Testen, Beurteilen, Bewerten* (pp. 44–55). Czwalina.
OECD. (2023). Main economic indicators. *OECDiLibrary*, *2023*.
Oelkers, L., Vogel, M., Kalenda, A., Surup, H. C., Korner, A., Kratzsch, J., & Kiess, W. (2021). Socioeconomic status is related to pubertal development in a german cohort. *Hormone Research in Paediatrics*, *93*. https://doi.org/10.1159/000513787
Ogden, C. L., & Flegal, K. M. (2010). Changes in terminology for childhood overweight and obesity. *National Health Statistics Reports*.
Olds, T. S. (2009). One million skinfolds: Secular trends in the fatness of young people 1951-2004. *European Journal of Clinical Nutrition*, *63*. https://doi.org/10.1038/ejcn.2009.7

Oliveira, M. S. R. D., Seabra, A., Freitas, D., Eisenmann, J. C., & Maia, J. (2014). Physical fitness percentile charts for children aged 6-10 from portugal. *The Journal of Sports Medicine and Physical Fitness*, *54*.

Onis, M. D., Onyango, A. W., Borghi, E., Siyam, A., Nishida, C., & Siekmann, J. (2007). Development of a WHO growth reference for school-aged children and adolescents. *Bulletin of the World Health Organization*, *85*. https://doi.org/10.2471/BLT.07.043497

Ortega, F. B., Artero, E. G., Ruiz, J. R., España-Romero, V., Jiménez-Pavón, D., Vicente-Rodriguez, G., Moreno, L. A., Manios, Y., Béghin, L., Ottevaere, C., Ciarapica, D., Sarri, K., Dietrich, S., Blair, S. N., Kersting, M., Molnar, D., González-Gross, M., Gutiérrez, Á., Sjöström, M., & Castillo, M. J. (2011). Physical fitness levels among european adolescents: The HELENA study. *British Journal of Sports Medicine*, *45*. https://doi.org/10.1136/bjsm.2009.062679

Ortega, F. B., Cadenas-Sánchez, C., Sánchez-Delgado, G., Mora-González, J., Martínez-Téllez, B., Artero, E. G., Castro-Piñero, J., Labayen, I., Chillón, P., Löf, M., & Ruiz, J. R. (2014). Systematic review and proposal of a field-based physical fitness-test battery in preschool children: The PREFIT battery. In *Sports Medicine* (Vol. 45). https://doi.org/10.1007/s40279-014-0281-8

Ortega, F. B., Leskošek, B., Blagus, R., Gil-Cosano, J. J., Mäestu, J., Tomkinson, G. R., Ruiz, J. R., Mäestu, E., Starc, G., Milanovic, I., Tammelin, T. H., Sorić, M., Scheuer, C., Carraro, A., Kaj, M., Csányi, T., Sardinha, L. B., Lenoir, M., Emeljanovas, A., ... Jurak, G. (2023). European fitness landscape for children and adolescents: Updated reference values, fitness maps and country rankings based on nearly 8 million test results from 34 countries gathered by the FitBack network. *British Journal of Sports Medicine*, *57*, 299–310. https://doi.org/10.1136/bjsports-2022-106176

Ortega, F. B., Ruiz, J. R., Castillo, M. J., & Sjöström, M. (2008). Physical fitness in childhood and adolescence: A powerful marker of health. In *International Journal of Obesity* (Vol. 32, pp. 1–11). Nature Publishing Group. https://doi.org/10.1038/sj.ijo.0803774

Oudshoorn, N. (2005). Beyond the natural body: An archaeology of sex hormones. In *Beyond the Natural Body: An Archaeology of Sex Hormones*. https://doi.org/10.4324/9780203421529

Pajek, S. V. (2022). Impact of the COVID-19 pandemic on the motor development of schoolchildren in rural and urban environments. *BioMed Research International*. https://doi.org/https://doi.org/10.1155/2022/8937693

Palomäki, S., Heikinaro-Johansson, P., & Huotari, P. (2015). Cardiorespiratory performance and physical activity in normal weight and overweight finnish adolescents from 2003 to 2010. *Journal of Sports Sciences*, *33*. https://doi.org/10.1080/02640414.2014.951874

Papadatou-Pastou, M., Ntolka, E., Schmitz, J., Martin, M., Munafò, M. R., Ocklenburg, S., & Paracchini, S. (2020). Human handedness: A meta-analysis. *Psychological Bulletin*, *146*. https://doi.org/10.1037/bul0000229

Park, M. H., Falconer, C., Viner, R. M., & Kinra, S. (2012). The impact of childhood obesity on morbidity and mortality in adulthood: A systematic review. *Obesity Reviews*, *13*. https://doi.org/10.1111/j.1467-789X.2012.01015.x

Pate, R. R. (1983). A new definition of youth fitness. *Physician and Sportsmedicine*, *11*. https://doi.org/10.1080/00913847.1983.11708509

Pate, R., Oria, M., & Pillsbury, L. (2012). Fitness measures and health outcomes in youth. In *Fitness Measures and Health Outcomes in Youth*. https://doi.org/10.17226/13483

Paterson, D. C., Ramage, K., Moore, S. A., Riazi, N., Tremblay, M. S., & Faulkner, G. (2021). Exploring the impact of COVID-19 on the movement behaviors of children and youth: A scoping review of evidence after the first year. In *Journal of Sport and Health Science* (Vol. 10). https://doi.org/10.1016/j.jshs.2021.07.001

Petermann, F., & Petermann, U. (2011). *Wechsler intelligence scale for children® – fourth edition - manual 1: Grundlagen, testauswertung und interpretation*. Pearson Assessment & Information GmbH.

Quek, Y. H., Tam, W. W. S., Zhang, M. W. B., & Ho, R. C. M. (2017). Exploring the association between childhood and adolescent obesity and depression: A meta-analysis. In *Obesity Reviews* (Vol. 18). https://doi.org/10.1111/obr.12535

R Core Team. (2022). *R: The r project for statistical computing* [Computer software].

Ramírez-Vélez, R., Tordecilla-Sanders, A., Correa-Bautista, J. E., Peterson, M. D., & Garcia-Hermoso, A. (2016). Handgrip strength and ideal cardiovascular health among colombian children and adolescents. *Journal of Pediatrics*, *179*. https://doi.org/10.1016/j.jpeds.2016.08.099

Raven, J., Raven, J. C., & H., C. J. (1998). *Standard progressive matrices*. Harcourt. Ravens-Sieberer, U., Ellert, U., & Erhart, M. (2007). Gesundheitsbezogene lebensqualität von kindern und jugendlichen in deutschland. *Bundesgesundheitsblatt - Gesundheitsforschung - Gesundheitsschutz*, *50*. https://doi.org/10.1007/s00103-007-0244-4

Ravindra, K., Malik, V. S., Padhi, B. K., Goel, S., & Gupta, M. (2022). Asymptomatic infection and transmission of COVID-19 among clusters: Systematic review and meta-analysis. In *Public Health* (Vol. 203). https://doi.org/10.1016/j.puhe.2021.12.003

Reed, J. A., Maslow, A. L., Long, S., & Hughey, M. (2013). Examining the impact of 45 minutes of daily physical education on cognitive ability, fitness performance, and body composition of african american youth. *Journal of Physical Activity and Health*, *10*. https://doi.org/10.1123/jpah.10.2.185

Reitan, R. (2004). The trail making test as an initial screening procedure for neuropsychological impairment in older children. *Archives of Clinical Neuropsychology*, *19*, 281–288. https://doi.org/10.1016/S0887-6177(03)00042-8

Reitan, R. M. (1971). Trail making test results for normal and brain-damaged children. *Perceptual and Motor Skills*, *33*, 575–581. https://doi.org/10.2466/pms.1971.33.2.575

Reitan, R. M., & Wolfson, D. (1995). Category test and trail making test as measures of frontal lobe functions. *The Clinical Neuropsychologist*, *9*. https://doi.org/10.1080/13854049508402057

Rittsteiger, L., Hinz, T., Oriwol, D., Wäsche, H., Santos-Hövener, C., & Woll, A. (2021). Sports participation of children and adolescents in germany: Disentangling the influence of parental socioeconomic status. *BMC Public Health*, *21*. https://doi.org/10.1186/s12889-021-11284-9

Rivera-Garza, C., & Herdt, G. (1996). Third sex, third gender: Beyond sexual dimorphism in culture and history. *Ethnohistory*, *43*. https://doi.org/10.2307/483350

Roberts, C. (2016). Tanner's puberty scale: Exploring the historical entanglements of children, scientific photography and sex. *Sexualities*, *19*. https://doi.org/10.1177/1363460715593477

Roth, K., & Willimczik, K. (1999). *Bewegungswissenschaft.* Rowohlt Taschenbuch,.
RStudio Team. (2022). RStudio: Integrated development for r. In *RStudio, Inc., Boston, MA* [Computer software].

Ruiz, J. R., Castro-Piñero, J., Artero, E. G., Ortega, F. B., Sjöström, M., Suni, J., & Castillo, M. J. (2009).

Predictive validity of health-related fitness in youth: A systematic review. In *British Journal of Sports Medicine* (Vol. 43). https://doi.org/10.1136/bjsm.2008.056499

Ruiz, J. R., Castro-Piñero, J., España-Romero, V., Artero, E. G., Ortega, F. B., Cuenca, M. A. M., Enez-Pavón, D. J., Chillón, P., Girela-Rejón, M. J., Mora, J., Gutiérrez, Á., Suni, J., Sjöström, M., & Castillo, M. J. (2011). Field-based fitness assessment in young people: The ALPHA health-related fitness test battery for children and adolescents. *British Journal of Sports Medicine*, *45*. https://doi.org/10.1136/bjsm.2010.075341

Ruiz-Ariza, A., Grao-Cruces, A., Loureiro, N. E. M. D., & Martï¿½nez-Lï¿½pez, E. J. (2017). Influence of physical fitness on cognitive and academic performance in adolescents: A systematic review from 2005–2015. *International Review of Sport and Exercise Psychology*, *10*. https://doi.org/10.1080/1750984X.2016.1184699

Safrit, M. J. (1990). The validity and reliability of fitness tests for children: A review. *Pediatric Exercise Science*, 9–28.

Salthouse, T. A. (1996). The processing-speed theory of adult age differences in cognition. *Psychological Review*, *103*. https://doi.org/10.1037/0033-295X.103.3.403

Santamaría-Fernández, P., & Fernández-Pinto, I. (2013). *Adaptación española de RIAS*. TEA Ediciones.

Santos, M. A. M. dos, Nevill, A. M., Buranarugsa, R., Pereira, S., Gomes, T. N. Q. F., Reyes, A., Barnett, L. M., & Maia, J. A. R. (2018). Modeling children's development in gross motor coordination reveals key modifiable determinants. An allometric approach. *Scandinavian Journal of Medicine and Science in Sports*, *28*. https://doi.org/10.1111/sms.13061

Santos, R., Mota, J., Santos, D. A., Silva, A. M., Baptista, F., & Sardinha, L. B. (2014). Physical fitness percentiles for portuguese children and adolescents aged 10-18 years. *Journal of Sports Sciences*, *32*. https://doi.org/10.1080/02640414.2014.906046

Sartorio, A., Lafortuna, C. L., Pogliaghi, S., & Trecate, L. (2002). The impact of gender, body dimension and body composition on hand-grip strength in healthy children. *Journal of Endocrinological Investigation*, *25*. https://doi.org/10.1007/BF03344033

Schilling, F., & Kiphard, E. J. (1974). *Körperkoordinationstest für kinder. KTK. Beltz; 1974.* KTK.

Schmidt, S. C. E., Anedda, B., Burchartz, A., Eichsteller, A., Kolb, S., Nigg, C., Niessner, C., Oriwol, D., Worth, A., & Woll, A. (2020). Physical activity and screen time of children and adolescents before and during the COVID-19 lockdown in germany: A natural experiment. *Scientific Reports*, *10*. https://doi.org/10.1038/s41598-020-78438-4

Schmidt, S. C. E., Burchartz, A., Kolb, S., Niessner, C., Oriwol, D., & Woll, A. (2022). Influence of socioeconomic variables on physical activity and screen time of children and adolescents during the COVID-19 lockdown in germany: The MoMo study. *German Journal of Exercise and Sport Research*, *52*, 362–373. https://doi.org/10.1007/s12662-021-00783-x

Schmidt, S., Will, N., Henn, A., Reimers, A., & Woll, A. (2016). *Der motorik-modul aktivitätsfragebogen MoMo-AFB - leitfaden zur anwendung und auswertung* (K. Bös, A. Worth, & A. Woll, Eds.). Karlsruher Institut für Technologie.

Schulz, S. (2013). *The reliability of the star co-ordination run and the 1-kg medicine ball push– physical fitness tests used in the EMOTIKON-study*. University Potsdam.

Seidman, S. (2009). Deconstructing queer theory or the under-theorization of the social and the ethical. In *Social Postmodernism*. https://doi.org/10.1017/cbo9780511520792.007

Seisdedos, N. (2012). *Adaptación española D2, test de atención de brickenkamp* (4th ed.). TEA Editociones.

Shallice, T. (1982). Specific impairments of planning. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *298*. https://doi.org/10.1098/rstb.1982.0082

Silva, S., Bustamante, A., Nevill, A., Katzmarzyk, P. T., Freitas, D., Prista, A., & Maia, J. (2016). An allometric modelling approach to identify the optimal body shape associated with, and differences between brazilian and peruvian youth motor performance. *PLoS ONE*, *11*. https://doi.org/10.1371/journal.pone.0149493

Simon, J. R., & Rudell, A. P. (1967). Auditory s-r compatibility: The effect of an irrelevant cue on information processing. *Journal of Applied Psychology*, *51*, 300–304. https://doi.org/10.1037/h0020586

Singh, A. S., Mulder, C., Twisk, J. W. R., Mechelen, W. V., & Chinapaw, M. J. M. (2008). Tracking of childhood overweight into adulthood: A systematic review of the literature. In *Obesity Reviews* (Vol. 9). https://doi.org/10.1111/j.1467-789X.2008.00475.x

Skog, H., Lintu, N., Haapala, H. L., & Haapala, E. A. (2020). Associations of cardiorespiratory fitness, adiposity, and arterial stiffness with cognition in youth. *Physiological Reports*, *8*. https://doi.org/10.14814/phy2.14586

Slaughter, M. H., Lohman, T. G., Boileau, R. A., Horswill, C. A., Stillman, R. J., Loan, M. D. V., & Bemben, D. A. (1988). Skinfold equations for estimations of body fatness in children and youth. *Human Biology*, *60*.

Smith, J. J., Eather, N., Morgan, P. J., Plotnikoff, R. C., Faigenbaum, A. D., & Lubans, D. R. (2014). The health benefits of muscular fitness for children and adolescents: A systematic review and meta-analysis. In *Sports Medicine* (Vol. 44, pp. 1209–1223). Springer International Publishing. https://doi.org/10.1007/s40279-014-0196-4

Sobel, V., Zhu, Y. S., & Imperato-McGinley, J. (2004). Fetal hormones and sexual differentiation. In *Obstetrics and Gynecology Clinics of North America* (Vol. 31). https://doi.org/10.1016/j.ogc.2004.08.005

Soriano, J. B., Murthy, S., Marshall, J. C., Relan, P., & Diaz, J. V. (2022). A clinical case definition of post-COVID-19 condition by a delphi consensus. In *The Lancet Infectious Diseases* (Vol. 22). https://doi.org/10.1016/S1473-3099(21)00703-9

Steinberger, J., Jacobs, D. R., Raatz, S., Moran, A., Hong, C. P., & Sinaiko, A. R. (2005). Comparison of body fatness measurements by BMI and skinfolds vs dual energy x-ray absorptiometry and

their relation to cardiovascular risk factors in adolescents. *International Journal of Obesity*, *29*, 1346– 1352. https://doi.org/10.1038/sj.ijo.0803026
Sullivan, D., & Hickel, J. (2023). Capitalism and extreme poverty: A global analysis of real wages, human height, and mortality since the long 16th century. *World Development*, *161*, 106026. https://doi.org/10.1016/j.worlddev.2022.106026
Tambalis, K. D., Panagiotakos, D. B., Psarra, G., Daskalakis, S., Kavouras, S. A., Geladas, N., Tokmakidis, S., & Sidossis, L. S. (2016). Physical fitness normative values for 6–18-year-old greek boys and girls, using the empirical distribution and the lambda, mu, and sigma statistical method. *European Journal of Sport Science*, *16*. https://doi.org/10.1080/17461391.2015.1088577
Tanner, J. M. (1951). Some notes on the reporting of growth data. *Human Biology; an International Record of Research*, *23*.
Tanner, J. M., & Whitehouse, R. H. (1981). *Atlas of children's growth - normal variation and growth disorders*. Academic Press Inc.
Taquet, M., Sillett, R., Zhu, L., Mendel, J., Camplisson, I., Dercon, Q., & Harrison, P. J. (2022). Neurological and psychiatric risk trajectories after SARS-CoV-2 infection: An analysis of 2-year retro-spective cohort studies including 1 284 437 patients. *Lancet Psychiatry*, *9*, 815–827.
Tate, S. A. (2009). Black beauty: Aesthetics, stylization, politics. In *Black Beauty: Aesthetics, Stylization, Politics*.
Teich, P. (2023, January). *EMOTIKON - grundschulsport*. https://www.uni--otsdam.de/de/emotikon/index
Teich, P., Fühner, T., Bähr, F., Puta, C., Granacher, U., & Kliegl, R. (2023). The covid pandemic affected the physical fitness of primary school children. *Sports Medicine - Open*, *9*.
Thompson, A., & Steinbeis, N. (2020). Sensitive periods in executive function development. In *Current Opinion in Behavioral Sciences* (Vol. 36). https://doi.org/10.1016/j.cobeha.2020.08.001
Todd, B. K., Barry, J. A., & Thommessen, S. A. O. (2017). Preferences for "gender-typed" toys in boys and girls aged 9 to 32 months. *Infant and Child Development*, *26*. https://doi.org/10.1002/icd.1986
Tomkinson, G. R. (2004). *Secular changes in fitness performance of australasian children and adolescents*. University of South Australia.
Tomkinson, G. R. (2007). Global changes in anaerobic fitness test performance of children and adolescents (1958-2003). *Scandinavian Journal of Medicine & Science in Sports*, *17*, 497–507. https://doi.org/10.1111/j.1600-0838.2006.00569.x
Tomkinson, G. R., Carver, K. D., Atkinson, F., Daniell, N. D., Lewis, L. K., Fitzgerald, J. S., Lang, J. J., & Ortega, F. B. (2018). European normative values for physical fitness in children and adolescents aged 9-17 years: Results from 2 779 165 eurofit performances representing 30 countries. *British Journal of Sports Medicine*, *52*. https://doi.org/10.1136/bjsports-2017-098253
Tomkinson, G. R., Kaster, T., Dooley, F. L., Fitzgerald, J. S., Annandale, M., Ferrar, K., Lang, J. J., & Smith, J. J. (2020). Temporal trends in the standing broad jump performance of 10,940,801 children and adolescents between 1960 and 2017. In *Sports Medicine* (Vol. 51). https://doi.org/10.1007/s40279-020-01394-6
Tomkinson, G. R., Lang, J. J., & Tremblay, M. S. (2019). Temporal trends in the cardiorespiratory fitness of children and adolescents representing 19 high-income and upper middle-income countries between 1981 and 2014. *British Journal of Sports Medicine*. https://doi.org/10.1136/bjsports-2017-097982
Tomkinson, G. R., Léger, L. A., Olds, T. S., & Cazorla, G. (2003). Secular trends in the performance of children and adolescents (1980-2000): An analysis of 55 studies of the 20m shuttle run test in 11 countries. In *Sports Medicine* (Vol. 33, pp. 285–300). https://doi.org/10.2165/00007256-200333040-00003
Tomkinson, G. R., & Olds, T. S. (2007). Secular changes in pediatric aerobic fitness test performance: The global picture. In *Pediatric Fitness* (Vol. 50, pp. 46–66). KARGER. https://doi.org/10.1159/000101075

Toombs, S. K. (2001). Handbook of phenomenology and medicine. *Philosophy and Medicine*. Towse, J. N., & McLachlan, A. (1999). An exploration of random generation among children. *British Journal of Developmental Psychology*, *17*. https://doi.org/10.1348/026151099165348

Tyndall, A. V., Clark, C. M., Anderson, T. J., Hogan, D. B., Hill, M. D., Longman, R. S., & Poulin, M. J. (2018). Protective effects of exercise on cognition and brain health in older adults. *Exercise and Sport Sciences Reviews*, *46*. https://doi.org/10.1249/JES.0000000000000161

Tyrrell, V. J., Richards, G., Hofman, P., Gillies, G. F., Robinson, E., & Cutfield, W. S. (2001). Foot-to-foot bioelectrical impedance analysis: A valuable tool for the measurement of body composition in children. *International Journal of Obesity*, *25*. https://doi.org/10.1038/sj.ijo.0801531

UNDP. (2018). Human development indices and indicators. 2018 statistical update. *United Nations Development Programme*, *27*.

USDA. (1990). *Dietary guidelines for americans*. US Department for Agriculture; US Department for Health; Human services.

Valdivia, A. B., Maia, J., & Nevill, A. (2015). Identifying the ideal body size and shape characteristics associated with children's physical performance tests in peru. *Scandinavian Journal of Medicine and Science in Sports*, *25*. https://doi.org/10.1111/sms.12231

Valle-Inclán, F. (1996). The locus of interference in the simon effect: An ERP study. *Biological Psychology*, *43*. https://doi.org/10.1016/0301-0511(95)05181-3

Venables, W. N., & Ripley, B. D. (2002). Modern applied statistics with s fourth edition by. In *World* (4th ed., Vol. 53).

Viciana, J., Mayorga-Vega, D., & Cocca, A. (2013). Effects of a maintenance resistance training program on muscular strength in schoolchildren. *Kinesiology: International Journal of Fundamental and Applied Kinesiology*, *45*, 82–91.

Wang, Y., & Lobstein, T. (2006). Worldwide trends in childhood overweight and obesity. *International Journal of Pediatric Obesity*, *1*. https://doi.org/10.1080/17477160600586747

Wang, Z., Zhang, J., Ying, Z., & Heymsfield, S. B. (2012). New insights into scaling of fat-free mass to height across children and adults. *American Journal of Human Biology*, *24*. https://doi.org/10.1002/ajhb.22286

Wechsler, D. (1991). *WISC-III: Wechsler intelligence scale for children*. Manual Psychological Corporation.

Wedderkopp, N., Jespersen, E., Franz, C., Klakk, H., Heidemann, M., Christiansen, C., Møller, N. C., & Leboeuf-Yde, C. (2012). Study protocol. The childhood health, activity, and motor performance school study denmark (the CHAMPS-study DK). *BMC Pediatrics*, *12* https://doi.org/10.1186/1471-2431-12-128

Welk, G. J., Morrow, J. R. J., & Falls, H. B. (2002). *FITNESSGRAM reference guide*. Cooper Institute.

Wells, N. M., Evans, G. W., Beavis, A., & Ong, A. D. (2010). Early childhood poverty, cumulative risk exposure, and body mass index trajectories through young adulthood. *American Journal of Public Health*, *100*. https://doi.org/10.2105/AJPH.2009.184291

Wendell, S. (1996). *The rejected body - feminist philosophical reflections on disability*. Routledge. Westbrook, L., & Schilt, K. (2014). Doing gender, determining gender: Transgender people, gender panics, and the maintenance of the sex/gender/sexuality system. *Gender and Society*, *28*. https://doi.org/10.1177/0891243213503203

Wickham, H. (2016). ggplot2 - elegant graphics for data analysis. In *Springer-Verlag New York* [Computer software].

Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, L., François, R., Grolemund, G., Hayes, A., Henry, L., Hester, J., Kuhn, M., Pedersen, T., Miller, E., Bache, S., Müller, K., Ooms, J., Robinson, D., Seidel, D., Spinu, V., ... Yutani, H. (2019). Welcome to the tidyverse. *Journal of Open Source Software*, *4*. https://doi.org/10.21105/joss.01686

Williams, D. P., Going, S. B., Lohman, T. G., Harsha, D. W., Srinivasan, S. R., Webber, L. S., & Berenson, G. S. (1992). Body fatness and risk for elevated blood pressure, total cholesterol, and serum lipoprotein ratios in children and adolescents. *American Journal of Public Health*, *82*. https://doi.org/10.2105/AJPH.82.3.358

Willis, K. (2020). Theories and practices of development. In *Theories and Practices of Development*. https://doi.org/10.4324/9781315559469

Wind, A. E., Takken, T., Helders, P. J. M., & Engelbert, R. H. H. (2010). Is grip strength a predictor for total muscle strength in healthy children, adolescents, and young adults? *European Journal of Pediatrics*, *169*. https://doi.org/10.1007/s00431-009-1010-4

Wohlfahrt-Veje, C., Tinggaard, J., Winther, K., Mouritsen, A., Hagen, C. P., Mieritz, M. G., Renzy-Martin, K. T. D., Boas, M., Petersen, J. H., & Main, K. M. (2014). Body fat throughout childhood in 2647 healthy danish children: Agreement of BMI, waist circumference, skinfolds with dual x-ray absorptiometry. *European Journal of Clinical Nutrition*, *68*. https://doi.org/10.1038/ejcn.2013.282

Wright, L., Steptoe, A., & Fancourt, D. (2020). Are we all in this together? Longitudinal assessment of cumulative adversities by socioeconomic position in the first 3 weeks of lockdown in the UK. *Journal of Epidemiology and Community Health*, *74*. https://doi.org/10.1136/jech-2020-214475

Xia, T., Song, L., Wang, T. T., Tan, L., & Mo, L. (2016). Exploring the effect of red and blue on cognitive task performances. *Frontiers in Psychology*, *7*. https://doi.org/10.3389/fpsyg.2016.00784

Xue, Y., Yang, Y., & Huang, T. (2019). Effects of chronic exercise interventions on executive function among children and adolescents: A systematic review with meta-analysis. In *British Journal of Sports Medicine* (Vol. 53). https://doi.org/10.1136/bjsports-2018-099825

Yamamoto, N., & Imai-Matsumura, K. (2019). Gender differences in executive function and behavioural self-regulation in 5 years old kindergarteners from east japan. *Early Child Development and Care*, *189*. https://doi.org/10.1080/03004430.2017.1299148

Yea, C., Barton, M., Bitnun, A., Morris, S. K., Tal, T. E., Ulloa-Gutierrez, R., Brenes-Chacon, H., Yock-Corrales, A., Ivankovich-Escoto, G., Soriano-Fallas, A., Mezerville, M. H., Gill, P., Nateghian, A., Aski, B. H., Manafi, A. A., Dwilow, R., Bullard, J., Papenburg, J., Scuccimarri, R., ... Yeh, E. A. (2023). Neurological involvement in hospitalized children with SARS-CoV-2 infection: A multinational study. *Canadian Journal of Neurological Sciences*. https://doi.org/10.1017/cjn.2022.347

Zimmermann, P., Pittet, L. F., & Curtis, N. (2021). How common is long COVID in children and adolescents? In *Pediatric Infectious Disease Journal* (Vol. 40). https://doi.org/10.1097/INF.0000000000003328

Part V

Appendix

8 R

8.1 Data preprocessing

The preprocessing of the data sets relevant for the analyses is documented here. The data sets are:

- *info* : contains Child- and School-IDs, Group (intervention allocation), as well as individual and assessment information
- *ages* : contains age and maturity offset information
- *anthro* : contains anthropometric measures body height, mass, body fat, and muscle mass
- *emo* : contains scores of EMOTIKON fitness tests (both scores and z-transformed scores)
- *cog* : contains scores of cognitive tests (both scores and z-transformed scores)

8.1.1 "info"

Used variables:

- *Child* : Participant ID (levels = "SMART01", "SMART02", ... , "SMART76") • School : School ID (levels = "01", "02", ... , "11")
- *Group* : Group affiliation (levels = "INT-CON", "CON-INT")
- *Time* : Assessment period (levels = "t0", "t1", "t2", "t3", "t6", "t8")
- *gender* : Participants gender (levels = "girl", "boy")
- *bp* : Berlin proximity of school (levels = "close", "far")
- *Date* : Mean date of assessment period
- *Dropout* : Information about dropouts at each assessment (levels = 0 [no], 1 [yes])

Participants were marked as "Dropout" according to the last assessment in which measures were collected.

8.1.2 "ages"

Used variables:

- *Child* : participant ID (levels = "SMART01", "SMART02", ... , "SMART76")
- *Time* : assessment period (levels = "t0", "t1", "t2", "t3", "t6", "t8")
- *age* : individual age at each assessment (years)
- *m_mirwald* : calculation of maturity offset according to Mirwald et al. 2002 (Mirwald et al., 2002) (years)

$$
\begin{aligned}
girls: -9.376 &+ (.0001882 * (stand - sit) * sit) \\
&+ (.0022 * (stand - sit) * age) \\
&+ (.005841 * age * sit) \\
&+ (-.002658 * age * weight) \\
&+ (.07693 * weight/stand * 100) \\
boys: -9.236 &+ (.0002708 * (stand - sit) * sit) \\
&+ (-.001663 * (stand - sit) * age) \\
&+ (.007216 * age * sit) \\
&+ (.02292 * weight/stand * 100)
\end{aligned}
$$

- *m_moore* : calculation of maturity offset according to Moore et al. 2012 (Moore et al., 2015) (years)

  $girls : -7.709133 + (.0042232 * (age * stand))$

  $boys : -8.128741 + (.0070346 * (age * sit))$

8.1.3 "anthro"

Used variables:

- *height* : body height (cm)
- *mass* : body mass (kg)
- *bmi* : body mass index (kg/m2)
- *fat* : body fat percent (%)
- *muscle* : muscle mass (kg)
- *lean* : lean mass (kg)

8.1.4 "emo"

Scores for the 20-m sprint and the star run where transformed from time to completion (s) into average speed (m/s) (Fühner et al., 2021, 2022). Z-scores were based norm values for key age children published in Fühner et al. 2021 (Fühner et al., 2021) for 6 min run, star run, 20 m sprint, standing long jump, and ball push test. As indicated by preliminary EMOTIKON analyses, one leg balance test were log-transformed and then z-transformed across the SMaRTER sample.

Table 8. 1: EMOTIKON values according to Fühner et al. 2021

Test	mean	sd
Run (m)	1004	148
Star_r (m/s)	2.05	0.288
S20_r (m/s)	4.52	0.413
SLJ (cm)	126	19.3
BPT (m)	3.74	714
OLB (s)	2.08	0.768

Run = 6 min run, Star_r = star run, S20_r = 20 m sprint, SLJ = standing long jump, BPT = ball push test, OLB = one leg balance

8.1.5 "cog"

A Box-Cox distribution analysis (Box & Cox, 1964) suggested logarithmic transformations for all executive function tests to obtain normally distributed model residuals (see Figure 8.1). Accordingly, lag transformed means and standard deviations were used for z-transformation (see Table 8.2). Further, the z-transformed scores for the trail making test and Simon task were reversed, so positive improvements of z-values indicate and improvement in time to completion or reaction time.

Table 8. 2: Mean and standard deviation of executive function log-scores

Test	mean	sd
TMTa (s)	3.12	0.328
TMTb (s)	3.83	0.370
DSST (n)	3.47	0.244
SC (ms)	6.97	0.228
SI (ms)	7.03	0.223

TMTa = trail making test verion A, TMTb = trail making test version B, DSST = digit symbol substitution test, SC = Simon task congruent condition, SI = Simon task incongruent condition

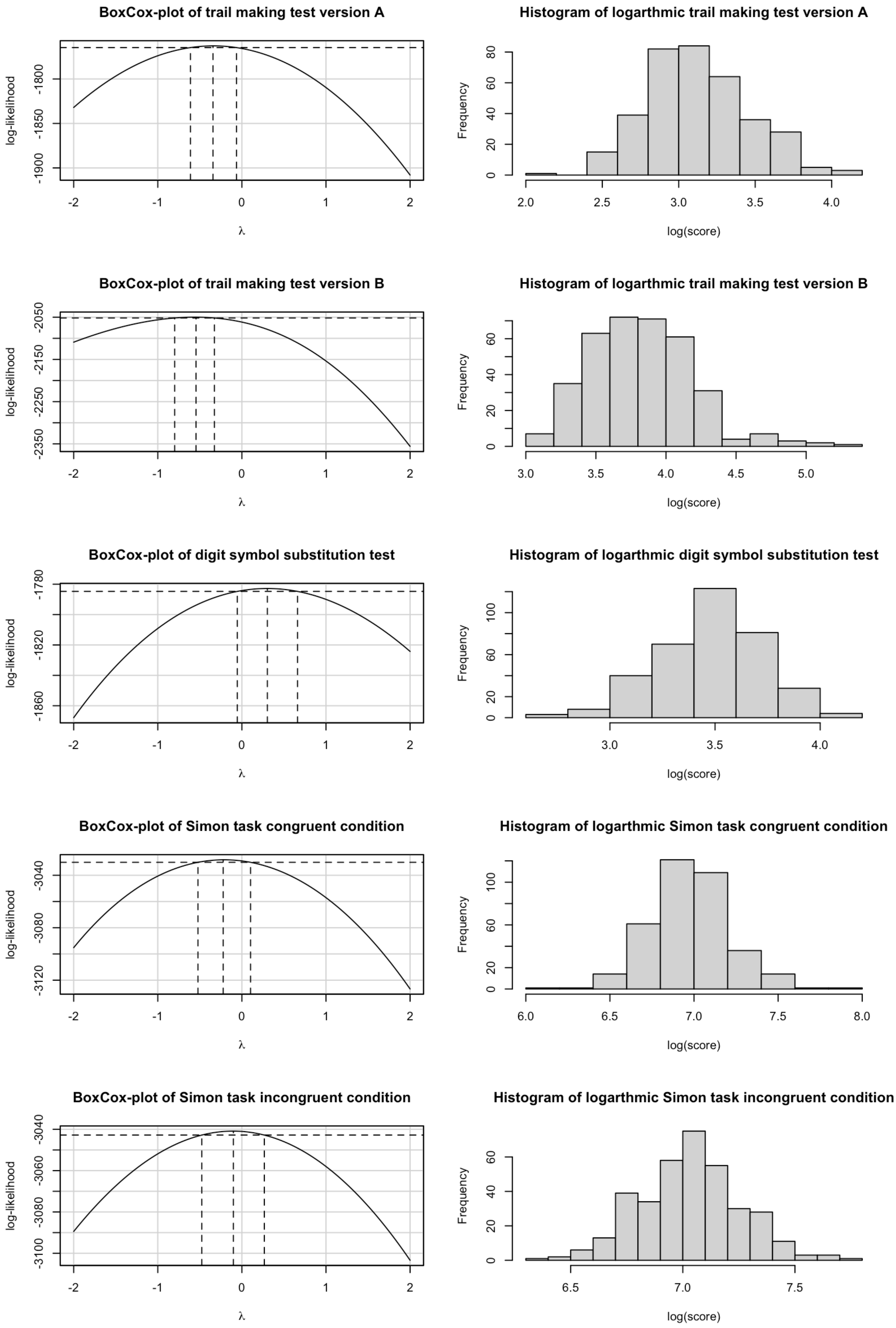


Figure 8. 1: Box-Cox distribution and histograms of executive function tests

8.2 Supplementary Material Chapter 2

8.2.1 Body mass index analysis for t1 and t3 populations

Table 8. 3: Group x gender x Time interactions for body mass index for the first intervention period (i.e., t0 to t1)

Predictors	Est.	SE	z	p
Grand mean	18.50	0.81	22.98	**<0.001**
Group2-1	0.62	1.61	0.38	0.703
gender2-1	-2.30	1.01	-2.28	**0.024**
Time2-1	0.22	0.07	3.21	**0.002**
Group2-1 * gender2-1	-3.94	2.02	-1.95	0.053
Group2-1 * Time2-1	0.02	0.14	0.12	0.905
gender2-1* Time2-1	0.19	0.14	1.37	0.172
Group2-1 * gender2-1 * Time2-1	0.28	0.28	0.99	0.324
Random Effects				
σ2	0.16			
τ00Child	16.28			
τ00school	4.49			
ICC	0.99			
N Child	76			
N school	11			
Observations	144			
Marginal R2 / Conditional R2	0.119 / 0.993			

Est. = estimate; SE = standar error

Table 8. 4 : Group x gender x Time interactions for body mass index for both intervention periods (i.e., t0 to t3)

Predictors	Est.	SE	z	p
Grand mean	18.76	0.82	22.95	**<0.001**
Group2-1	0.39	1.63	0.24	0.812
gender2-1	-2.20	1.02	-2.16	**0.032**
Time2-1	0.20	0.11	1.82	0.070
Time3-2	0.28	0.11	2.50	**0.013**
Time4-3	0.42	0.12	3.48	**0.001**
Group2-1 * gender2-1	-4.25	2.04	-2.08	**0.039**
Group2-1 * Time2-1	0.01	0.22	0.05	0.962
Group2-1 * Time3-2	-0.42	0.22	-1.86	0.064
Group2-1 * Time4-3	-0.14	0.24	-0.59	0.553
gender2-1 * Time2-1	0.20	0.22	0.91	0.365
gender2-1 * Time3-2	0.11	0.22	0.50	0.615
gender2-1 * Time4-3	0.23	0.24	0.95	0.345
Group2-1 * gender2-1 * Time2-1	0.17	0.43	0.40	0.693
Group2-1 * gender2-1 * Time3-2	0.12	0.45	0.27	0.791
Group2-1 * gender2-1 * Time4-3	-1.59	0.48	-3.32	**0.001**
Random Effects				
σ2	0.38			
τ00Child	16.62			
τ00school	4.64			
ICC	0.98			
N Child	76			
N school	11			
Observations	268			
Marginal R2 / Conditional R2	0.120 / 0.984			

Est. = estimate; SE = standar error

8.3 Supplementary Material Chapter 3

8.3.1 INT M1: physical fitness

8.3.1.1 Model fitting

Table 8. 5: LRT 1 of fixed effects in INT M1 - physical fitness

model	npar	AIC	BIC	logLik	deviance	Chisq	Df	p	sig
m1.1_emo	6	1,227.027	1,253.005	-607.513	1,215.027				
m1.2_emo	18	1,232.915	1,310.850	-598.457	1,196.915	18.112	12	0.112	
m1.3_emo_abp	21	1,228.853	1,319.777	-593.427	1,186.853	10.061	3	0.018	*
m1.4_emo_all	22	1,230.662	1,325.916	-593.331	1,186.662	0.191	1	0.662	
m1.1_emo	6	1,227.027	1,253.005	-607.513	1,215.027				
m1.2_emo	18	1,232.915	1,310.850	-598.457	1,196.915	18.112	12	0.112	
m1.3_emo_gbp	21	1,229.324	1,320.248	-593.662	1,187.324	9.591	3	0.022	*
m1.4_emo_all	22	1,230.662	1,325.916	-593.331	1,186.662	0.661	1	0.416	
m1.1_emo	6	1,227.027	1,253.005	-607.513	1,215.027				
m1.2_emo	18	1,232.915	1,310.850	-598.457	1,196.915	18.112	12	0.112	
m1.3_emo_gap	21	1,234.319	1,325.243	-596.159	1,192.319	4.596	3	0.204	
m1.4_emo_all	22	1,230.662	1,325.916	-593.331	1,186.662	5.657	1	0.017	*
m1.1_emo	6	1,227.027	1,253.005	-607.513	1,215.027				
m1.2_emo	18	1,232.915	1,310.850	-598.457	1,196.915	18.112	12	0.112	
m1.3_emo_gab	21	1,229.686	1,320.610	-593.843	1,187.686	9.229	3	0.026	*
m1.4_emo_all	22	1,230.662	1,325.916	-593.331	1,186.662	1.023	1	0.312	

npar = number of parameters; AIC = Akaike information criterion; BIC = Bayesian information criterion; loglik; Chisp = chi square; DF = degrees of freedom

LRT indicated that body mass index (BMI) improves the model's fit, while age, gender, and Berlin proximity can be dropped.

Table 8. 6: LRT 2 of fixed effects in INT M1 - physical fitness

Model	npar	AIC	BIC	logLik	deviance	Chisq	Df	p	sig
m1.1_emo	6	1,227.027	1,253.005	-607.513	1,215.027				
m1.2_emo	18	1,232.915	1,310.850	-598.457	1,196.915	18.112	12	0.112	
m1.5_emo_bmi	19	1,226.565	1,308.830	-594.282	1,188.565	8.350	1	0.004	**
m1.6_emo_bmi	22	1,185.591	1,280.845	-570.796	1,141.591	46.973	3	0.000	***
m1.7_emo_bmi	30	1,195.350	1,325.242	-567.675	1,135.350	6.241	8	0.620	

npar = number of parameters; AIC = Akaike information criterion; BIC = Bayesian information criterion; loglik; Chisp = chi square; DF = degrees of freedom

LRT further showed that including BMI nested within Test without interactions improves the models fit. Accordingly, m1.6_emo_bmi was reported in Section 3.3.1.1.

8.3.1.2 Means and standard deviations of model variables

Table 8. 7: Means and standard deviation of the model variables used in INT M1 physical fitness

Time	t0		t1	
Group	CON	INT	CON	INT
gender [girl/boy]	16/15	16/25	16/14	14/25
Berlin proximity [close/far]	15/16	0/41	15/15	0/39
BMI [kg/m²]	18.4±4.7	19.1±5	18.5±4.8	19.5±4.9
age [years]	9.1±0.5	9.2±0.6	9.4±0.5	9.5±0.6
20 m sprint [m/s]	4.1±0.4	4.1±0.3	4.1±0.5	4.1±0.3
standing long jump [cm]	107.1±23.9	106.4±13.6	110.2±24.7	106.8±14.2
star run [m/s]	1.8±0.2	1.8±0.2	1.9±0.2	1.8±0.2
6 min run [m]	876.8±129.9	834.1±119.3	896.1±144.3	863.5±125.1

CON = control condition; INT = intervention condition; BMI = body mass index

8.3.2 INT M1: executive function

8.3.2.1 Model fitting

Table 8. 8: LRT 1 of fixed effects in INT M1 - executive function

Model	npar	AIC	BIC	logLik	deviance	Chisq	Df	p	sig
m1.1_cog	7	1,776.276	1,808.183	-881.138	1,762.276				
m1.2_cog	22	1,765.195	1,865.475	-860.597	1,721.195	41.081	15	0.000	***
m1.3_cog_abp	25	1,769.455	1,883.410	-859.727	1,719.455	1.740	3	0.628	
m1.4_cog_all	26	1,764.845	1,883.359	-856.423	1,712.845	6.609	1	0.010	*
m1.1_cog	7	1,776.276	1,808.183	-881.138	1,762.276				
m1.2_cog	22	1,765.195	1,865.475	-860.597	1,721.195	41.081	15	0.000	***
m1.3_cog_gbp	25	1,762.846	1,876.801	-856.423	1,712.846	8.349	3	0.039	*
m1.4_cog_all	26	1,764.845	1,883.359	-856.423	1,712.845	0.001	1	0.982	
m1.1_cog	7	1,776.276	1,808.183	-881.138	1,762.276				
m1.2_cog	22	1,765.195	1,865.475	-860.597	1,721.195	41.081	15	0.000	***
m1.3_cog_gap	25	1,763.515	1,877.470	-856.758	1,713.515	7.679	3	0.053	
m1.4_cog_all	26	1,764.845	1,883.359	-856.423	1,712.845	0.670	1	0.413	
m1.1_cog	7	1,776.276	1,808.183	-881.138	1,762.276				
m1.2_cog	22	1,765.195	1,865.475	-860.597	1,721.195	41.081	15	0.000	***
m1.3_cog_gab	25	1,763.611	1,877.566	-856.806	1,713.611	7.583	3	0.055	
m1.4_cog_all	26	1,764.845	1,883.359	-856.423	1,712.845	0.766	1	0.381	

npar = number of parameters; AIC = Akaike information criterion; BIC = Bayesian information criterion; loglik; Chisp = chi square; DF = degrees of freedom

LRT indicated that gender improves the model's fit, while age, BMI, and Berlin proximity can be dropped.

Table 8. 9: LRT 2 of fixed effects in INT M1 - executive function

Model	npar	AIC	BIC	logLik	deviance	Chisq	Df	p	sig
m1.1_cog	7	1,776.276	1,808.183	-881.138	1,762.276				
m1.2_cog	22	1,765.195	1,865.475	-860.597	1,721.195	41.081	15	0.000	***
m1.5_cog_gender	23	1,761.022	1,865.861	-857.511	1,715.022	6.172	1	0.013	*
m1.6_cog_gender	27	1,761.152	1,884.223	-853.576	1,707.152	7.871	4	0.096	
m1.7_cog_gender	37	1,772.997	1,941.650	-849.498	1,698.997	8.155	10	0.614	

npar = number of parameters; AIC = Akaike information criterion; BIC = Bayesian information criterion; loglik; Chisp = chi square; DF = degrees of freedom

LRT further showed that gender added as a fixed factor not nested within Test improves the models fit. Accordingly, m1.5_cog_bmi was reported in Section 3.3.2.1.

8.3.2.2 Means and standard deviations of model variables

Table 8. 10: Means and standard deviation of the model variables used in INT M1 executive function

Time	t0		t1	
Group	CON	INT	CON	INT
gender [girl/boy]	16/15	16/25	16/14	14/25
Berlin proximity [close/far]	15/16	0/41	15/15	0/39
BMI [kg/m²]	18.4±4.7	19.1±5	18.5±4.8	19.5±4.9
age [years]	9.1±0.5	9.2±0.6	9.4±0.5	9.5±0.6
TMT version A [s]	30.3±8.3	31.2±11.1	26.5±8.4	25.1±7.4
TMT version B [s]	64.7±34.4	59.3±19.6	58.6±22.8	53.1±20.3
DSST [n]	28.5±5.2	27.1±5.3	30.4±6.5	30.6±6.2
SC [ms]	1198.1±202.5	1275.3±340.3	1191.3±273.8	1185.2±283.6
SI [ms]	1277.7±217.5	1335.1±279.1	1205.7±190.6	1244.9±301.8

CON = control condition; INT = intervention condition; TMT = trail making test; DSST = digit symbol substitution test; SC = Simon task congruent condition; Simon task incongruent condition; BMI = body mass index

8.3.3 INT M2: physical fitness

8.3.3.1 Model fitting

Table 8. 11: LRT 1 of fixed effects in INT M2 - physical fitness

Model	npar	AIC	BIC	logLik	deviance	Chisq	Df	p	Sig
m2.1_emo	6	2,176.077	2,205.636	-1,082.038	2,164.077				
m2.2_emo	18	2,176.732	2,265.411	-1,070.366	2,140.732	23.344	12	0.025	*
m2.3_emo_abp	21	2,168.538	2,271.996	-1,063.269	2,126.538	14.194	3	0.003	**
m2.4_emo_all	22	2,169.783	2,278.167	-1,062.891	2,125.783	0.756	1	0.385	
m2.1_emo	6	2,176.077	2,205.636	-1,082.038	2,164.077				
m2.2_emo	18	2,176.732	2,265.411	-1,070.366	2,140.732	23.344	12	0.025	*
m2.3_emo_gbp	21	2,171.247	2,274.705	-1,064.623	2,129.247	11.486	3	0.009	**
m2.4_emo_all	22	2,169.783	2,278.167	-1,062.891	2,125.783	3.464	1	0.063	
m2.1_emo	6	2,176.077	2,205.636	-1,082.038	2,164.077				
m2.2_emo	18	2,176.732	2,265.411	-1,070.366	2,140.732	23.344	12	0.025	*
m2.3_emo_gap	21	2,175.276	2,278.734	-1,066.638	2,133.276	7.456	3	0.059	
m2.4_emo_all	22	2,169.783	2,278.167	-1,062.891	2,125.783	7.494	1	0.006	**
m2.1_emo	6	2,176.077	2,205.636	-1,082.038	2,164.077				
m2.2_emo	18	2,176.732	2,265.411	-1,070.366	2,140.732	23.344	12	0.025	*
m2.3_emo_gab	21	2,170.199	2,273.657	-1,064.100	2,128.199	12.533	3	0.006	**
m2.4_emo_all	22	2,169.783	2,278.167	-1,062.891	2,125.783	2.417	1	0.120	

npar = number of parameters; AIC = Akaike information criterion; BIC = Bayesian information criterion; loglik; Chisp = chi square; DF = degrees of freedom

LRT indicated that BMI improves the model's fit, while age, gender, and Berlin proximity can be dropped.

Table 8. 12: LRT 2 of fixed effects in INT M2 - physical fitness

Model	npar	AIC	BIC	logLik	deviance	Chisq	Df	p	sig
m2.1_emo	6	2,176.077	2,205.636	-1,082.038	2,164.077				
m2.2_emo	18	2,176.732	2,265.411	-1,070.366	2,140.732	23.344	12	0.025	*
m2.5_emo_bmi	19	2,170.194	2,263.799	-1,066.097	2,132.194	8.538	1	0.003	**
m2.6_emo_bmi	22	2,113.439	2,221.824	-1,034.720	2,069.439	62.755	3	0.000	***
m2.7_emo_bmi	30	2,123.595	2,271.392	-1,031.797	2,063.595	5.845	8	0.665	

npar = number of parameters; AIC = Akaike information criterion; BIC = Bayesian information criterion; loglik; Chisp = chi square; DF = degrees of freedom

LRT further showed that BMI added as a fixed factor nested within Test without interactions improves the models fit. Accordingly, m2.6_emo_bmi was reported in Section 3.4.1.1.

8.3.3.2 Means and standard deviations of model variables

Table 8. 13: Means and standard deviation of the model variables used in INT M2 physical fitness

Time [pooled]	pre		post	
Group [pooled]	CON	INT	CON	INT
gender [girl/boy]	28/38	27/38	28/36	26/36
Berlin proximity [close/far]	14/52	14/51	14/50	13/49
BMI [kg/m²]	19.1±4.8	19.1±4.9	19.3±5.1	19.8±4.9
age [years]	9.4±0.6	9.4±0.6	9.8±0.7	9.7±0.6
20 m sprint [m/s]	4.1±0.4	4.1±0.4	4.1±0.4	4.1±0.4
standing long jump [cm]	108.4±20.6	107.6±18.5	112.2±20.7	108.3±19.5
star run [m/s]	1.8±0.2	1.8±0.2	1.9±0.2	1.9±0.2
6 min run [m]	850.4±127	853.6±144	880.1±140	872.7±132.4

CON = control condition; INT = intervention condition; BMI = body mass index

8.3.4 INT M2: executive function

8.3.4.1 Model fitting

Table 8. 14: LRT 1 of fixed effects in INT M2 - executive function

Model	npar	AIC	BIC	logLik	deviance	Chisq	Df	p	sig
m2.1_cog	7	3,279.784	3,315.915	-1,632.892	3,265.784				
m2.2_cog	22	3,246.800	3,360.356	-1,601.400	3,202.800	62.984	15	0.000	***
m2.3_cog_abp	25	3,105.015	3,234.055	-1,527.507	3,055.015	147.785	3	0.000	***
m2.4_cog_all	26	3,106.509	3,240.712	-1,527.255	3,054.509	0.505	1	0.477	
m2.1_cog	7	3,279.784	3,315.915	-1,632.892	3,265.784				
m2.2_cog	22	3,246.800	3,360.356	-1,601.400	3,202.800	62.984	15	0.000	***
m2.3_cog_gbp	25	3,236.158	3,365.198	-1,593.079	3,186.158	16.642	3	0.001	***
m2.4_cog_all	26	3,106.509	3,240.712	-1,527.255	3,054.509	131.648	1	0.000	***
m2.1_cog	7	3,279.784	3,315.915	-1,632.892	3,265.784				
m2.2_cog	22	3,246.800	3,360.356	-1,601.400	3,202.800	62.984	15	0.000	***
m2.3_cog_gap	25	3,106.976	3,236.016	-1,528.488	3,056.976	145.824	3	0.000	***
m2.4_cog_all	26	3,106.509	3,240.712	-1,527.255	3,054.509	2.467	1	0.116	
m2.1_cog	7	3,279.784	3,315.915	-1,632.892	3,265.784				
m2.2_cog	22	3,246.800	3,360.356	-1,601.400	3,202.800	62.984	15	0.000	***
m2.3_cog_gab	25	3,104.689	3,233.730	-1,527.345	3,054.689	148.111	3	0.000	***
m2.4_cog_all	26	3,106.509	3,240.712	-1,527.255	3,054.509	0.180	1	0.671	

npar = number of parameters; AIC = Akaike information criterion; BIC = Bayesian information criterion; loglik; Chisp = chi square; DF = degrees of freedom

LRT indicated that age improves the model’s fit, while BMI, gender, and Berlin proximity can be dropped.

Table 8. 15: LRT 2 of fixed effects in INT M2 - executive function

Model	npar	AIC	BIC	logLik	deviance	Chisq	Df	p	sig
m2.1_cog	7	3,279.784	3,315.915	-1,632.892	3,265.784				
m2.2_cog	22	3,246.800	3,360.356	-1,601.400	3,202.800	62.984	15	0.000	***
m2.5_cog_gender	23	3,103.466	3,222.184	-1,528.733	3,057.466	145.334	1	0.000	***
m2.6_cog_gender	27	3,106.360	3,245.724	-1,526.180	3,052.360	5.107	4	0.277	
m2.7_cog_gender	37	3,114.710	3,305.690	-1,520.355	3,040.710	11.650	10	0.309	

npar = number of parameters; AIC = Akaike information criterion; BIC = Bayesian information criterion; loglik; Chisp = chi square; DF = degrees of freedom

LRT further showed that age added as a fixed factor not nested within Test improves the models fit. Accordingly, m2.5_cog_bmi was reported in Section 3.4.2.1.

8.3.4.2 Means and standard deviations of model variables

Table 8. 16: Means and standard deviation of the model variables used in INT M2 executive function

Time [pooled]	pre		post	
Group [pooled]	CON	INT	CON	INT
gender [girl/boy]	28/38	27/38	28/37	26/37
Berlin proximity [close/far]	14/52	14/51	14/51	14/49
BMI [kg/m²]	19.1±4.8	19.1±4.9	19.3±5	19.7±4.9
age [years]	9.4±0.6	9.4±0.6	9.8±0.7	9.7±0.6
TMT version A [s]	25.8±8.3	27.9±10.5	23.9±7.3	23.6±7.4
TMT version B [s]	55.4±27.5	59.1±24.5	46.7±19.1	47.8±17.5
DSST [n]	31.4±6.6	29.4±6.9	33.3±6.9	32.3±6.8
SC [ms]	1133.1±224.9	1168.9±319	1104.6±225.6	1106.5±273.6
SI [ms]	1209.4±236.8	1249.8±270	1133±187.7	1176.2±303.1

CON = control condition; INT = intervention condition; TMT = trail making test; DSST = digit symbol substitution test; SC = Simon task congruent condition; Simon task incongruent condition; BMI = body mass index

8.3.5 INT M3: physical fitness

8.3.5.1 Model fitting

Table 8. 17: LRT 1 of fixed effects in INT M3 - physical fitness

Model	npar	AIC	BIC	logLik	deviance	Chisq	Df	p	sig
m3.1_emo	6	1,918.185	1,946.556	-953.092	1,906.185				
m3.2_emo	50	1,873.974	2,110.405	-886.987	1,773.974	132.211	44	0.000	***
m3.3_emo_abp	53	1,862.963	2,113.580	-878.482	1,756.963	17.011	3	0.001	***
m3.4_emo_all	54	1,864.902	2,120.248	-878.451	1,756.902	0.061	1	0.804	
m3.1_emo	6	1,918.185	1,946.556	-953.092	1,906.185				
m3.2_emo	50	1,873.974	2,110.405	-886.987	1,773.974	132.211	44	0.000	***
m3.3_emo_gbp	53	1,862.951	2,113.568	-878.475	1,756.951	17.023	3	0.001	***
m3.4_emo_all	54	1,864.902	2,120.248	-878.451	1,756.902	0.049	1	0.824	
m3.1_emo	6	1,918.185	1,946.556	-953.092	1,906.185				
m3.2_emo	50	1,873.974	2,110.405	-886.987	1,773.974	132.211	44	0.000	***
m3.3_emo_gap	53	1,870.924	2,121.541	-882.462	1,764.924	9.050	3	0.029	*
m3.4_emo_all	54	1,864.902	2,120.248	-878.451	1,756.902	8.022	1	0.005	**
m3.1_emo	6	1,918.185	1,946.556	-953.092	1,906.185				
m3.2_emo	50	1,873.974	2,110.405	-886.987	1,773.974	132.211	44	0.000	***
m3.3_emo_gab	53	1,867.524	2,118.142	-880.762	1,761.524	12.450	3	0.006	**
m3.4_emo_all	54	1,864.902	2,120.248	-878.451	1,756.902	4.623	1	0.032	*

npar = number of parameters; AIC = Akaike information criterion; BIC = Bayesian information criterion; loglik; Chisp = chi square; DF = degrees of freedom

LRT indicated that BMI and Berlin proximity improves the model's fit, while age and gender can be dropped.

Table 8. 18: LRT 2 of fixed effects in INT M3 - physical fitness

Model	npar	AIC	BIC	logLik	deviance	Chisq	Df	p	sig
m3.1_emo	6	1,918.185	1,946.556	-953.092	1,906.185				
m3.2_emo	50	1,873.974	2,110.405	-886.987	1,773.974	132.211	44	0.000	***
m3.5_emo_bmi	51	1,863.558	2,104.718	-880.779	1,761.558	12.416	1	0.000	***
m3.6_emo_bmi	54	1,795.105	2,050.451	-843.553	1,687.105	74.453	3	0.000	***
m3.7_emo_bmi	78	1,827.327	2,196.160	-835.664	1,671.327	15.778	24	0.896	
m3.1_emo	6	1,918.185	1,946.556	-953.092	1,906.185				
m3.2_emo	50	1,873.974	2,110.405	-886.987	1,773.974	132.211	44	0.000	***
m3.5_emo_bp	51	1,867.127	2,108.287	-882.564	1,765.127	8.847	1	0.003	**
m3.6_emo_bp	54	1,840.301	2,095.647	-866.151	1,732.301	32.826	3	0.000	***
m3.7_emo_bp	74	1,855.488	2,205.406	-853.744	1,707.488	24.813	20	0.209	

npar = number of parameters; AIC = Akaike information criterion; BIC = Bayesian information criterion; loglik; Chisp = chi square; DF = degrees of freedom

LRT indicated that BMI and Berlin proximity added in the fixed structure nested within Test improves the model's fit.

Table 8. 19: LRT 3 of fixed effects in INT M3 - physical fitness

Model	npar	AIC	BIC	logLik	deviance	Chisq	Df	p	sig
m3.1_emo	6	1,918.185	1,946.556	-953.092	1,906.185				
m3.2_emo	50	1,873.974	2,110.405	-886.987	1,773.974	132.211	44	0.000	***
m3.6_emo_bmi	54	1,795.105	2,050.451	-843.553	1,687.105	86.869	4	0.000	***
m3.6_emo_bmibp	58	1,788.050	2,062.310	-836.025	1,672.050	15.055	4	0.005	**
m3.6_emo_bmibp2	62	1,789.523	2,082.698	-832.762	1,665.523	6.527	4	0.163	
m3.1_emo	6	1,918.185	1,946.556	-953.092	1,906.185				
m3.2_emo	50	1,873.974	2,110.405	-886.987	1,773.974	132.211	44	0.000	***
m3.6_emo_bp	54	1,840.301	2,095.647	-866.151	1,732.301	41.673	4	0.000	***
m3.6_emo_bmibp	58	1,788.050	2,062.310	-836.025	1,672.050	60.251	4	0.000	***
m3.6_emo_bmibp2	62	1,789.523	2,082.698	-832.762	1,665.523	6.527	4	0.163	

npar = number of parameters; AIC = Akaike information criterion; BIC = Bayesian information criterion; loglik; Chisp = chi square; DF = degrees of freedom

LRT further indicated that BMI and Berlin proximity both added to the fixed structure nested within Test without interactions improves the model's fit. Accordingly, m3.6_emo_bmibp was reported in Section 3.5.1.1.

8.3.5.2 Means and standard deviations of model variables

Table 8. 20: Means and standard deviation of the model variables used in INT M3 physical fitness

Time	t0		t1		t2		t3		t6		t8	
Group	CON-INT	INT-CON	CON-INT	INT-CON	CON-INT	INT-CON	CON-INT	INT-CON	CON-INT	INT-CON	CON-INT	INT-CON
gender [girl/boy]	9/9	7/12	9/8	6/12	9/9	7/12	9/7	7/12	7/7	7/10	9/8	7/12
Berlin proximity [close/far]	8/10	0/19	8/9	0/18	8/10	0/19	7/9	0/19	8/6	0/17	7/10	0/19
BMI [kg/m²]	19±5	20±6	19±5	21±6	19±5	20±6	20±6	21±6	19±6	20±5	21±6	22±7
age [years]	9±0.4	9.2±0.5	9.3±0.4	9.5±0.5	9.5±0.4	9.7±0.5	9.9±0.4	10.1±0.5	10.5±0.4	10.6±0.4	11.5±0.4	11.7±0.5
20 m sprint [m/s]	4.1±0.4	4±0.4	4.1±0.5	4±0.4	4.1±0.3	4.2±0.3	4±0.4	4.1±0.3	4.2±0.4	4.3±0.3	4.3±0.4	4.4±0.5
standing long jump [cm]	111±21	107±14	112±24	106±16	109±18	107±16	109±20	116±18	114±29	120±18	122±26	126±16
star run [m/s]	1.8±0.2	1.7±0.2	1.9±0.2	1.8±0.2	1.8±0.2	1.8±0.2	1.9±0.3	1.9±0.2	2±0.2	2±0.2	2±0.2	2±0.3
6 min run [m]	879±138	794±146	865±133	837±156	854±169	791±103	856±137	837±169	881±196	835±110	825±185	860±174

CON = control condition; INT = intervention condition; BMI = body mass index

8.3.6 INT M3: executive function

8.3.6.1 Model fitting

Table 8. 21: LRT 1 of fixed effects in INT M3 - executive function

Model	npar	AIC	BIC	logLik	deviance	Chisq	Df	p	sig
m3.1_cog	7	2,832.854	2,867.629	-1,409.427	2,818.854				
m3.2_cog	62	2,451.983	2,759.994	-1,163.992	2,327.983	490.870	55	0.000	***
m3.3_cog_abp	65	2,455.093	2,778.007	-1,162.547	2,325.093	2.890	3	0.409	
m3.4_cog_all	66	2,454.411	2,782.293	-1,161.206	2,322.411	2.682	1	0.102	
m3.1_cog	7	2,832.854	2,867.629	-1,409.427	2,818.854				
m3.2_cog	62	2,451.983	2,759.994	-1,163.992	2,327.983	490.870	55	0.000	***
m3.3_cog_gbp	65	2,452.707	2,775.621	-1,161.354	2,322.707	5.276	3	0.153	
m3.4_cog_all	66	2,454.411	2,782.293	-1,161.206	2,322.411	0.296	1	0.587	
m3.1_cog	7	2,832.854	2,867.629	-1,409.427	2,818.854				
m3.2_cog	62	2,451.983	2,759.994	-1,163.992	2,327.983	490.870	55	0.000	***
m3.3_cog_gap	65	2,454.039	2,776.953	-1,162.020	2,324.039	3.944	3	0.268	
m3.4_cog_all	66	2,454.411	2,782.293	-1,161.206	2,322.411	1.628	1	0.202	
m3.1_cog	7	2,832.854	2,867.629	-1,409.427	2,818.854				
m3.2_cog	62	2,451.983	2,759.994	-1,163.992	2,327.983	490.870	55	0.000	***
m3.3_cog_gab	65	2,452.597	2,775.511	-1,161.299	2,322.597	5.386	3	0.146	
m3.4_cog_all	66	2,454.411	2,782.293	-1,161.206	2,322.411	0.186	1	0.666	

npar = number of parameters; AIC = Akaike information criterion; BIC = Bayesian information criterion; loglik; Chisp = chi square; DF = degrees of freedom

LRT indicated that including BMI, Berlin proximity, age, and gender in the fixed factors do not improve the model's fit. Accordingly, all covariates were dropped and m3.2_cog was reported in Section 3.5.2.1.

8.3.6.2 Means and standard deviations of model variables

Table 8. 22: Means and standard deviation of the model variables used in INT M3 executive function

Time	t0		t1		t2		t3		t6		t8	
Group	CON-INT	INT-CON	CON-INT	INT-CON	CON-INT	INT-CON	CON-INT	INT-CON	CON-INT	INT-CON	CON-INT	INT-CON
gender [girl/boy]	9/9	7/12	9/8	6/12	9/9	7/12	9/8	7/12	7/7	7/10	9/9	7/12
Berlin proximity [close/far]	8/10	0/19	8/9	0/18	8/10	0/19	8/9	0/19	8/6	0/17	8/10	0/19
BMI [kg/m²]	19±5	20±6	19±5	21±6	19±5	20±6	20±6	21±6	19±6	20±5	21±6	22±7
age [years]	9±0.4	9.2±0.5	9.3±0.4	9.5±0.5	9.5±0.4	9.7±0.5	9.9±0.4	10.1±0.5	10.5±0.4	10.6±0.4	11.5±0.4	11.7±0.5
TMT version A [s]	28.8±8.6	31.1±11.2	24.5±7.3	24.7±7	22.3±6.7	22.7±7.6	22.3±7.7	23±8.1	19.8±4.9	19.4±5.9	15.5±3.9	18.3±4.4
TMT version B [s]	57.1±18.9	60.5±18.9	54.6±28	52.5±24.9	57.9±36.3	43.4±12.9	39.9±9.2	39.3±8.1	37.8±11.3	40.4±12.1	33.6±9.9	34.4±9
DSST [n]	29.9±5.9	27.1±5.9	31.5±6.8	29.7±6.3	33.4±8.1	33.7±7.7	36.2±7.8	35.3±7	37.9±6.3	35.4±9.1	45.2±8	40.7±7
SC [ms]	1157±200	1286±412	1139±223	1170±300	1026±250	1070±259	943±156	1056±174	985±203	992±180	845±97	869±188
SI [ms]	1246±227	1335±283	1155±152	1232±296	1120±227	1158±272	1028±176	1112±199	1067±218	1089±273	888±133	903±184

CON = control condition; INT = intervention condition; TMT = trail making test; DSST = digit symbol substitution test; SC = Simon task congruent condition; Simon task incongruent condition; BMI = body mass index

8.4 Supplementary Material Chapter 4

8.4.1 Correlation of InBody parameters

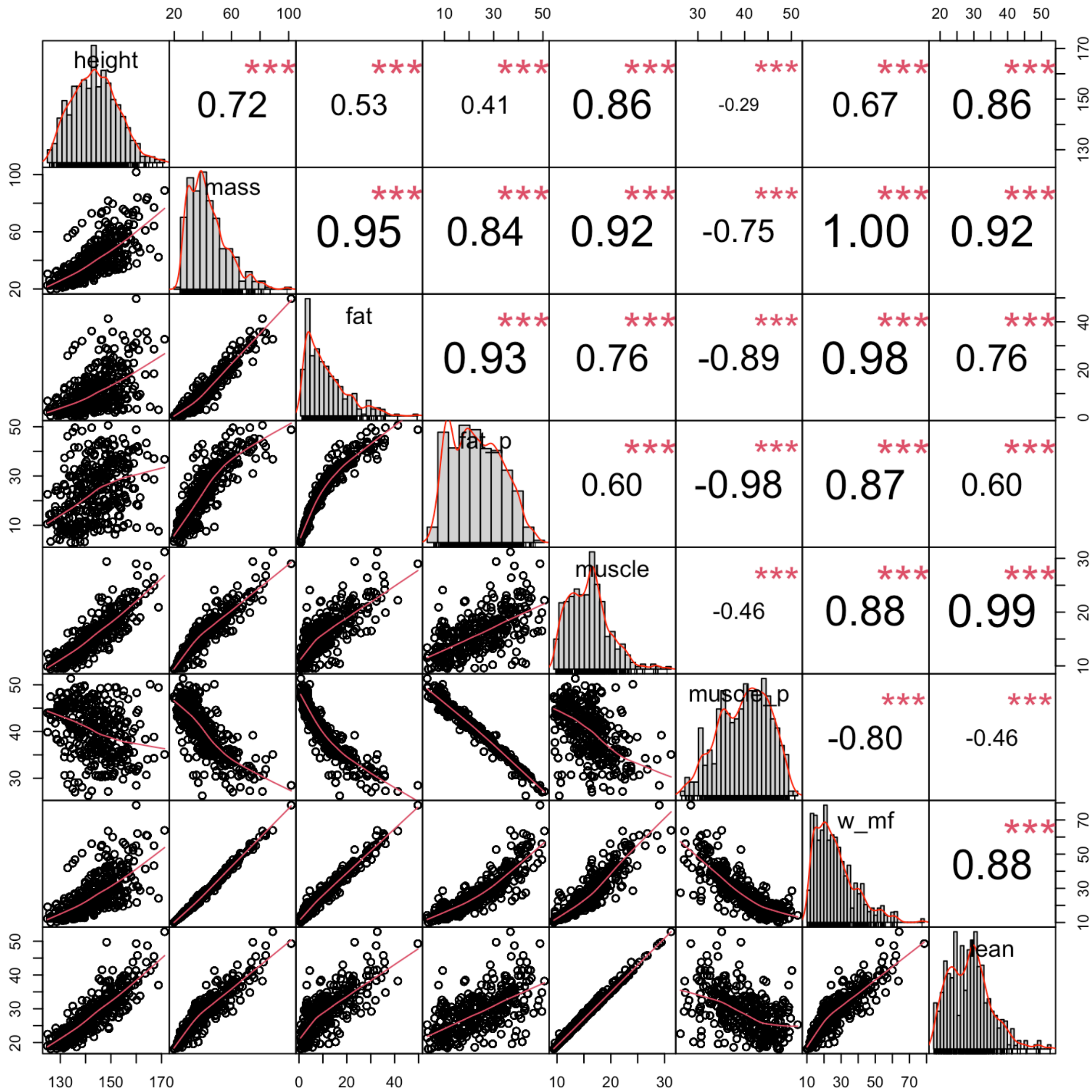


Figure 8. 2: Correlations for anthropometric parameters assessed using the InBody; height = body height [cm], mass = body mass [kg], fat = body fat mass [kg], fat_p = body fat percent [%], muscle = body muscle mass [kg], muscle_p = body muscle mass percent [%], w_mf = sum of muscle mass and fat mass [kg], lean = body lean mass

8.4.2 Model selection

Table 8. 23: LRT 1 of fixed effects

Model	npar	AIC	BIC	logLik	deviance	Chisq	Df	p	sig
m1_0	10	4,233.194	4,287.987	-2,106.597	4,213.194				
m1_m	15	3,644.544	3,726.734	-1,807.272	3,614.544	598.650	5	0	***
m1_hm	20	3,627.413	3,736.999	-1,793.706	3,587.413	27.131	5	0	***
m1_0	10	4,233.194	4,287.987	-2,106.597	4,213.194				
m1_h	15	3,938.921	4,021.110	-1,954.460	3,908.921	304.274	5	0	***
m1_hm	20	3,627.413	3,736.999	-1,793.706	3,587.413	321.508	5	0	***

npar = number of parameters; AIC = Akaike information criterion; BIC = Bayesian information criterion; loglik; Chisp = chi square; DF = degrees of freedom

Including body mass and height into the model, nested within Test, improve the fit.

Table 8. 24: LRT 2 of fixed effects

Model	npar	AIC	BIC	logLik	deviance	Chisq	Df	p	sig
m1_hm	20	3,627.413	3,736.999	-1,793.706	3,587.413				
m2_h2	25	3,629.874	3,766.856	-1,789.937	3,579.874	7.539	5	0.184	
m2_h2m2	30	3,621.639	3,786.018	-1,780.819	3,561.639	18.235	5	0.003	**
m1_hm	20	3,627.413	3,736.999	-1,793.706	3,587.413				
m2_m2	25	3,621.002	3,757.984	-1,785.501	3,571.002	16.411	5	0.006	**
m2_h2m2	30	3,621.639	3,786.018	-1,780.819	3,561.639	9.363	5	0.095	

npar = number of parameters; AIC = Akaike information criterion; BIC = Bayesian information criterion; loglik; Chisp = chi square; DF = degrees of freedom

LRT revealed only significant quadratic effects for body mass improves models fit.

Table 8. 25: LRT 3 of fixed effects

model	npar	AIC	BIC	logLik	deviance	Chisq	Df	p	sig
m2_m2	25	3,621.002	3,757.984	-1,785.501	3,571.002				
m3_hm1	30	3,615.165	3,779.544	-1,777.582	3,555.165	15.837	5	0.007	**
m3_hm2	35	3,621.845	3,813.620	-1,775.922	3,551.845	3.320	5	0.651	

npar = number of parameters; AIC = Akaike information criterion; BIC = Bayesian information criterion; loglik; Chisp = chi square; DF = degrees of freedom

Second-order interactions showed significant improvements in models fitted for interactions of body size with linear but not quadratic trends in body mass. Accordingly, m3_hm1 was reported in Section 4.2.

8.4.3 Means and standard deviations of model variables

Table 8. 26: Means and standard deviation of the model variables

Time	t0	t1	t2	t3	t6	t8
gender [girl/boy]	35/41	31/39	28/40	28/35	18/26	16/21
body height [cm]	138.6±7.8	140.9±7.8	142.1±7.8	144.6±8.5	148±7.7	154.1±8
body mass [kg]	36.4±12.3	38.4±12.7	39.5±12.8	42±14.2	44.7±13.4	51.9±18.7
age [years]	9.2±0.5	9.4±0.5	9.6±0.5	10.1±0.5	10.7±0.5	11.6±0.5
20 m sprint [m/s]	4.1±0.4	4.1±0.4	4.1±0.4	4.1±0.4	4.2±0.4	4.3±0.4
standing long jump [cm]	107.6±18.7	108.2±19.3	109.8±19.8	111.7±20.4	114±21	123.8±20.9
star run [m/s]	1.8±0.2	1.8±0.2	1.8±0.2	1.9±0.2	2±0.2	2±0.2
6 min run [m]	858.6±125.5	877.9±133.9	846.4±139.4	868±136.3	864.2±146.2	843.9±177.4
ball push test [m]	3.6±0.7	3.7±0.7	3.6±0.6	3.8±0.7	4.1±0.7	4.8±0.8

8.5 Supplementary Material Chapter 5

8.5.1 Correlation age, maturity, and their principal components

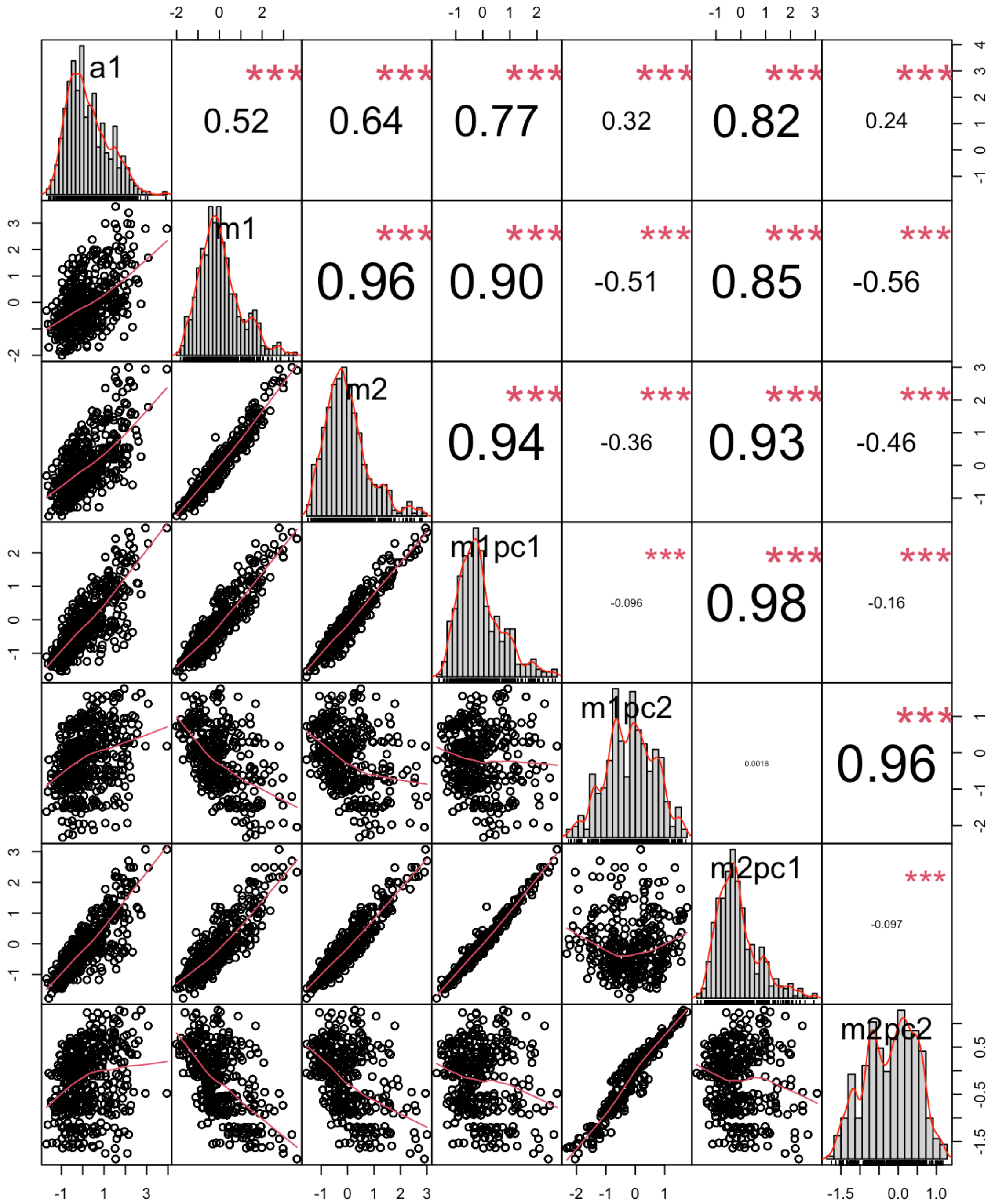


Figure 8. 3: Correlations for age-related parameters; a1 = age centered at 0 [years], m1 = maturity offset centered at 0 (Mirwald) [years], m2 = maturity offset centered at 0 (Moore) [years], m1pc1 = first principal component of age and maturity (Mirwald) [z-score], m1pc2 = second principal component of age and maturity (Mirwald) [z-score], m2pc1 = first principal component of age and maturity (Moore) [z-score], m2pc2 = second principal component of age and maturity (Moore) [z-score]

8.5.2 Model 1: Effect of maturity offset according to Mirwald, age, and gender on physical fitness

8.5.2.1 Model selection

Table 8. 27: LRT 1 of fixed effects in Model 1

Model	npar	AIC	BIC	logLik	deviance	Chisq	Df	p	sig
m1	2	3,251.739	3,262.246	-1,623.870	3,247.739				
m2_p1	14	3,093.607	3,167.155	-1,532.803	3,065.607	182.132	12	0	***
m2_all	18	3,005.397	3,099.959	-1,484.699	2,969.397	96.210	4	0	***
m1	2	3,251.739	3,262.246	-1,623.870	3,247.739				
m2_p2	14	3,021.588	3,095.137	-1,496.794	2,993.588	254.151	12	0	***
m2_all	18	3,005.397	3,099.959	-1,484.699	2,969.397	24.191	4	0	***
m1	2	3,251.739	3,262.246	-1,623.870	3,247.739				
m2_ge	14	3,019.143	3,092.691	-1,495.571	2,991.143	256.596	12	0	***
m2_all	18	3,005.397	3,099.959	-1,484.699	2,969.397	21.746	4	0	***

npar = number of parameters; AIC = Akaike information criterion; BIC = Bayesian information criterion; loglik; Chisp = chi square; DF = degrees of freedom

LRTs indicate that both principal components and gender improve the models fit.

Table 8. 28: LRT 2 of fixed effects in Model 1

model	npar	AIC	BIC	logLik	deviance	Chisq	Df	p	sig
m2_all	18	3,005.397	3,099.959	-1,484.699	2,969.397				
m311	22	3,010.470	3,126.046	-1,483.235	2,966.470	2.927	4	0.570	
m32	26	3,016.290	3,152.880	-1,482.145	2,964.290	2.180	4	0.703	
m33	34	3,012.921	3,191.539	-1,472.460	2,944.921	19.369	8	0.013	*
m2_all	18	3,005.397	3,099.959	-1,484.699	2,969.397				
m312	22	3,011.091	3,126.668	-1,483.546	2,967.091	2.306	4	0.680	
m32	26	3,016.290	3,152.880	-1,482.145	2,964.290	2.801	4	0.592	
m33	34	3,012.921	3,191.539	-1,472.460	2,944.921	19.369	8	0.013	*

npar = number of parameters; AIC = Akaike information criterion; BIC = Bayesian information criterion; loglik; Chisp = chi square; DF = degrees of freedom

LRTs reveal a significance third order interaction, however the model is very likely to be overparametrise according to AIC and BIC.

Table 8. 29: LRT 3 of fixed effects in Model 1

model	npar	AIC	BIC	logLik	deviance	Chisq	Df	p	sig
m2_all	18	3,005.397	3,099.959	-1,484.699	2,969.397				
m3_p1	20	2,987.174	3,092.243	-1,473.587	2,947.174	22.223	2	0	***
m3	23	2,967.072	3,087.902	-1,460.536	2,921.072	26.101	3	0	***
m2_all	18	3,005.397	3,099.959	-1,484.699	2,969.397				
m3_p2	20	2,990.027	3,095.096	-1,475.013	2,950.027	19.370	2	0	***
m3	23	2,967.072	3,087.902	-1,460.536	2,921.072	28.954	3	0	***

npar = number of parameters; AIC = Akaike information criterion; BIC = Bayesian information criterion; loglik; Chisp = chi square; DF = degrees of freedom

Both principal components entered into the random effect structure of Child did significantly improve the fit of the model. Accordingly, m3 will be reported in Section 5.3.

8.5.3 Model 2: Effect of maturity offset according to Moore, age, and gender on physical fitness

8.5.3.1 Model selection

Table 8. 30: LRT 1 of fixed effects in Model 2

Model	npar	AIC	BIC	logLik	deviance	Chisq	Df	p	sig
m4	2	3,251.739	3,262.246	-1,623.870	3,247.739				
m5_p1	14	3,133.981	3,207.529	-1,552.990	3,105.981	141.759	12	0.000	***
m5_all	18	3,025.249	3,119.812	-1,494.625	2,989.249	116.731	4	0.000	***
m4	2	3,251.739	3,262.246	-1,623.870	3,247.739				
m5_p2	14	3,029.002	3,102.551	-1,500.501	3,001.002	246.737	12	0.000	***
m5_all	18	3,025.249	3,119.812	-1,494.625	2,989.249	11.753	4	0.019	*
m4	2	3,251.739	3,262.246	-1,623.870	3,247.739				
m5_ge	14	3,030.911	3,104.460	-1,501.456	3,002.911	244.828	12	0.000	***
m5_all	18	3,025.249	3,119.812	-1,494.625	2,989.249	13.662	4	0.008	**

npar = number of parameters; AIC = Akaike information criterion; BIC = Bayesian information criterion; loglik; Chisp = chi square; DF = degrees of freedom

LRTs indicate that both principal components and gender improve the models fit.

Table 8. 31: LRT 2 of fixed effects in Model 2

model	npar	AIC	BIC	logLik	deviance	Chisq	Df	p	sig
m5_all	18	3,025.249	3,119.812	-1,494.625	2,989.249				
m511	22	3,031.705	3,147.282	-1,493.853	2,987.705	1.544	4	0.819	
m52	26	3,036.331	3,172.921	-1,492.165	2,984.331	3.374	4	0.497	
m53	34	3,032.803	3,211.421	-1,482.401	2,964.803	19.528	8	0.012	*
m5_all	18	3,025.249	3,119.812	-1,494.625	2,989.249				
m512	22	3,030.231	3,145.807	-1,493.115	2,986.231	3.019	4	0.555	
m52	26	3,036.331	3,172.921	-1,492.165	2,984.331	1.900	4	0.754	
m53	34	3,032.803	3,211.421	-1,482.401	2,964.803	19.528	8	0.012	*

npar = number of parameters; AIC = Akaike information criterion; BIC = Bayesian information criterion; loglik; Chisp = chi square; DF = degrees of freedom

LRTs reveal a significance third order interaction, however the model is very likely to be overparametrise according to AIC and BIC.

Table 8. 32: LRT 3 of fixed effects in Model 2

model	npar	AIC	BIC	logLik	deviance	Chisq	Df	p	sig
m5_all	18	3,025.249	3,119.812	-1,494.625	2,989.249				
m6_p1	20	3,008.201	3,113.270	-1,484.100	2,968.201	21.049	2	0	***
m6	23	2,990.853	3,111.683	-1,472.427	2,944.853	23.348	3	0	***
m5_all	18	3,025.249	3,119.812	-1,494.625	2,989.249				
m6_p2	20	3,007.803	3,112.872	-1,483.901	2,967.803	21.446	2	0	***
m6	23	2,990.853	3,111.683	-1,472.427	2,944.853	22.950	3	0	***

npar = number of parameters; AIC = Akaike information criterion; BIC = Bayesian information criterion; loglik; Chisp = chi square; DF = degrees of freedom

Both principal components entered into the random effect structure of Child did significantly improve the fit of the model. Accordingly, m6 will be reported in Section 5.4.

8.5.4 Means and standard deviations of model variables

Table 8. 33: Means and standard deviation of the model variables used in INT M1 physical fitness

Time	t0	t1	t2	t3	t6	t8
gender [girl/boy]	35/41	29/38	28/37	28/33	18/26	15/19
maturity offset (Mirwald) [years]	-2.9±0.8	-2.6±0.8	-2.5±0.8	-2.1±0.9	-1.7±0.9	-0.8±1
maturity offset (Moore) [years]	-2.9±0.6	-2.7±0.6	-2.6±0.6	-2.2±0.7	-1.7±0.8	-0.9±0.8
age [years]	9.2±0.5	9.4±0.5	9.6±0.5	10.1±0.6	10.7±0.5	11.6±0.5
20 m sprint [m/s]	4.1±0.4	4.1±0.4	4.1±0.4	4.1±0.4	4.2±0.4	4.3±0.4
standing long jump [cm]	107.6±18.7	108±19.6	109.8±20.1	111.9±20.7	114±21	123.6±21.3
star run [m/s]	1.8±0.2	1.8±0.2	1.8±0.2	1.9±0.2	2±0.2	2±0.3
6 min run [m]	858.6±125.5	878.8±134.7	846.4±139.4	868±136.3	864.2±146.2	843.9±177.4

9 Julia

9.1 Supplementary Material Chapter 6

9.1.1 Main effects Simon Task 9.1.1.1 Model selection

- m0: speed ~ 1 + Congruence + Colour + Side + a1 + gender + (1 | Child)
- m0C1: speed ~ 1 + Congruence * a1 + gender + Colour + Side + (1 | Child)
- m0C2: speed ~ 1 + Congruence * (a1 + gender) + Colour + Side + (1 | Child)

No improvements of the models fit due to interactions between the Congruence with age and gender.

- m1: speed ~ 1 + Congruence + Side + Colour + a1 + gender + (1 + Congruence + a1 | Child)
- m2: speed ~ 1 + Congruence + Side + Colour + a1 + gender + (1 + Congruence + Side + Colour + a1 | Child)

Including Congruence, Colour, Side, and age as variance components in the random effect structure of Child improves the models fit.

- m3: speed ~ 1 + Congruence + Side + Colour + a1 + gender + zerocorr(1 + Congruence + Side + Colour + a1 | Child)

Keeping all correlation parameters from the random effect structure of Child has the better fit.

- m4: speed ~ 1 + Congruence + Side + Colour + a1 + gender + (1 + a1 | Child) + zerocorr(0 + Congruence + Side + Colour | Child)

Table 9. 1: LRT of m0, m0C1, and m0C2

Row	name	dof	deviance	AIC	AICc	BIC	χ^2	χ^2_dof	pvalue
1	m0	8	882.824	898.824	898.831	962.41			
2	m0C1	9	881.472	899.472	899.481	971.007	1.351	1.0	0.245
3	m0C2	10	880.731	900.731	900.742	980.214	0.741	1.0	0.389

Table 9. 2: LRT of m0, m1, and m2

Row	name	dof	deviance	AIC	AICc	BIC	χ^2	χ^2_dof	pvalue
1	m0	8	882.824	898.824	898.831	962.41			
2	m1	13	500.56	526.56	526.577	629.887	382.264	5.0	0.0
3	m2	22	381.669	425.669	425.717	600.531	118.891	9.0	0.0

Table 9. 3: LRT of m2 and m3

Row	name	dof	deviance	AIC	AICc	BIC	χ^2	χ^2_dof	pvalue
1	m3	12	408.198	432.198	432.212	527.577			
2	m2	22	381.669	425.669	425.717	600.531	26.529	10.0	0.003

Table 9. 4: LRT of m2, m4 and m3

Row	name	dof	deviance	AIC	AICc	BIC	χ^2	χ^2_dof	pvalue
1	m3	12	408.198	432.198	432.212	527.577			
2	m4	13	403.936	429.936	429.954	533.264	4.261	1.0	0.039
3	m2	22	381.669	425.669	425.717	600.531	22.267	9.0	0.008

Including the individual correlation parameters Congruence, Colour, Side, and age gives the best fit. Accordingly, m2 is reported in Section 6.3.2.

9.1.1.2 Means and standard deviations of model variabels

Table 9. 5: Means and standard deviations of Simon task reaction times by condition and assessment

Row	Condition	t0	t1	t2	t3	t6	t8
1	number of reaction times [n]	4388	4059	3904	3838	2586	2141
2	age [years]	9.2±0.5	9.4±0.5	9.6±0.5	10.1±0.5	10.7±0.5	11.6±0.5
3	gender [girls/boys]	35/41	30/39	28/40	28/37	18/26	16/21
4	blue left (incongruent)	1318±455	1206±427	1172±501	1090±434	1088±478	917±299
5	blue right (congruent)	1265±573	1200±575	1093±350	1047±425	1013±494	859±317
6	red left (congruent)	1231±505	1168±514	1092±514	1018±425	1049±499	897±398
7	red right (incongruent)	1303±481	1279±635	1204±679	1079±442	1119±524	896±290

9.1.2 Simon Task x EMOTIKON

9.1.2.1 Model selection

A minimal model m4_m0 was fitted according to the structure of m2. All EMOTIKON tests were included in m4_mall. Six additional models were fitted (m4_mRUN, m4_mStar, m4_mS20, m4_mSLJ, m4_mBPT, & m4_mOLB), with each dropping one EMOTIKON test.

- m4_m0: speed ~ 1 + Congruence + Side + Colour + a1 + gender + (1 + Congruence + Side + Colour + a1 | Child)
- m4_mRun: speed ~ 1 + Congruence + Side + Colour + a1 + gender + Star_r + S20_r + SLJ + BPT + OLB_l + (1 + Congruence + Side + Colour + a1 | Child)
- m4_mStar: speed ~ 1 + Congruence + Side + Colour + a1 + gender + Run + S20_r + SLJ + BPT + OLB_l + (1 + Congruence + Side + Colour + a1 | Child)
- m4_mS20: speed ~ 1 + Congruence + Side + Colour + a1 + gender + Run + Star_r + SLJ + BPT + OLB_l + (1 + Congruence + Side + Colour + a1 | Child)
- m4_mSLJ: speed ~ 1 + Congruence + Side + Colour + a1 + gender + Run + Star_r + S20_r + BPT + OLB_l + (1 + Congruence + Side + Colour + a1 | Child)

- m4_mBPT: speed ~ 1 + Congruence + Side + Colour + a1 + gender + Run + Star_r + S20_r + SLJ + OLB_l + (1 + Congruence + Side + Colour + a1 | Child)
- m4_mOLB: speed ~ 1 + Congruence + Side + Colour + a1 + gender + Run + Star_r + S20_r + SLJ + BPT + (1 + Congruence + Side + Colour + a1 | Child)
- m4_mall: speed ~ 1 + Congruence + Side + Colour + a1 + gender + Run + Star_r + S20_r + SLJ + BPT + OLB_l + (1 + Congruence + Side + Colour + a1 | Child)

Table 9. 6: LRT of m4_m0, m4_mRun, and m4_mall

Row	name	dof	deviance	AIC	AICc	BIC	χ^2	χ^2_dof	pvalue
1	m4_m0	22	410.105	454.105	454.155	628.218			
2	m4_mRun	27	367.127	421.127	421.202	634.811	42.978	5.0	0.0
3	m4_mall	28	371.32	427.32	427.401	648.919	-4.193	1.0	NaN

LRT reveals that including Run improves the models fit.

Table 9. 7: LRT of m4_m0, m4_mStar, and m4_mall

Row	name	dof	deviance	AIC	AICc	BIC	χ^2	χ^2_dof	pvalue
1	m4_m0	22	410.105	454.105	454.155	628.218			
2	m4_mStar	27	380.002	434.002	434.077	647.686	30.103	5.0	0.0
3	m4_mall	28	371.32	427.32	427.401	648.919	8.682	1.0	0.003

LRT reveals that including Star improves the models fit.

Table 9. 8: LRT of m4_m0, m4_mS20, and m4_mall

Row	name	dof	deviance	AIC	AICc	BIC	χ^2	χ^2_dof	pvalue
1	m4_m0	22	410.105	454.105	454.155	628.218			
2	m4_mS20	27	364.421	418.421	418.496	632.105	45.684	5.0	0.0
3	m4_mall	28	371.32	427.32	427.401	648.919	-6.899	1.0	NaN

LRT reveals that including S20_r does not improve the models fit.

Table 9. 9: LRT of m4_m0, m4_mSLJ, and m4_mall

Row	name	dof	deviance	AIC	AICc	BIC	χ^2	χ^2_dof	pvalue
1	m4_m0	22	410.105	454.105	454.155	628.218			
2	m4_mSLJ	27	361.089	415.089	415.164	628.773	49.016	5.0	0.0
3	m4_mall	28	371.32	427.32	427.401	648.919	-10.231	1.0	NaN

LRT reveals that including SLJ does not improve the models fit.

Table 9. 10: LRT of m4_m0, m4_mBPT, and m4_mall

Row	name	dof	deviance	AIC	AICc	BIC	χ^2	χ^2_dof	pvalue
1	m4_m0	22	410.105	454.105	454.155	628.218			
2	m4_mBPT	27	361.124	415.124	415.199	628.808	48.981	5.0	0.0
3	m4_mall	28	371.32	427.32	427.401	648.919	-10.196	1.0	NaN

LRT reveals that including BPT does not improve the models fit.

Table 9. 11: LRT of m4_m0, m4_mOLB, and m4_mall

Row	name	dof	deviance	AIC	AICc	BIC	χ²	χ²_dof	pvalue
1	m4_m0	22	410.105	454.105	454.155	628.218			
2	m4_mOLB	27	381.117	435.117	435.192	648.801	28.988	5.0	0.0
3	m4_mall	28	371.32	427.32	427.401	648.919	9.797	1.0	0.002

LRT reveals that including OLB improves the models fit.

Accordingly, a optimal model m5 using Run, Star, and OBL was fitted.

- m5: speed ~ 1 + Congruence + Side + Colour + a1 + gender + Run + Star_r + OLB_l + (1 + a1 + Congruence + Side + Colour | Child)

Table 9. 12: LRT of m4_m0, m5, and m4_mall

Row	name	dof	deviance	AIC	AICc	BIC	χ²	χ²_dof	pvalue
1	m4_m0	22	410.105	454.105	454.155	628.218			
2	m5	25	364.552	414.552	414.617	612.408	45.553	3.0	0.0
3	m4_mall	28	371.32	427.32	427.401	648.919	-6.768	3.0	NaN

Verification that dropping S20_r, SLJ, and BPT does not affect the models overall fit.

Checking for interactions of fixed effects:

- m6: speed ~ 1 + (Congruence + Side + Colour) * (Run + Star_r + OLB_l) + a1 + gender + (1 + a1 + Congruence + Side + Colour | Child)

Table 9. 13: LRT of m5, and m6

Row	name	dof	deviance	AIC	AICc	BIC	χ²	χ²_dof	pvalue
1	m5	25	364.552	414.552	414.617	612.408			
2	m6	34	358.744	426.744	426.862	695.828	5.808	9.0	0.759

LRT shows no improvement of model fit for including interactions between Simon task parameters (i.e., Congruence, Colour, and Side) and EMOTIKON tests (Run, Star, and OLB).

Next, Run, Star and, OLB are included in the random effects without CP (i.e., m7) and with (i.e., m8).

- m7: speed ~ 1 + Congruence + Side + Colour + a1 + gender + Run + Star_r + OLB_l + (1 + a1+ Congruence + Side + Colour | Child) + zerocorr(0 + Run + Star_r + OLB_l | Child)
- m8: speed ~ 1 + Congruence + Side + Colour + a1 + gender + Run + Star_r + OLB_l + (1 + a1+ Congruence + Side + Colour + Run + Star_r + OLB_l | Child)

Table 9. 14: LRT of m5, m7, and m8

Row	name	dof	deviance	AIC	AICc	BIC	χ^2	χ^2_dof	pvalue
1	m5	25	364.552	414.552	414.617	612.408			
2	m7	28	-350.44	-294.44	-294.36	-72.8415	714.992	3.0	0.0
3	m8	46	-414.837	-322.837	-322.623	41.2173	64.397	18.0	0.0

Including Run, Star, and OLB as variance components with associated correlation parameters in the random effects structure has the best fit.

Accordingly, m8 was reported in Section 6.4.2.

9.1.2.2 Fixed and random effects of m5

Table 9. 15: Fixed effects of m5

Name	Est.	SE	z	p
Grand mean	0.987	0.018	54.5	<0.001
Congruence: con	0.029	0.003	11.2	<0.001
Side: right	-0.004	0.002	-1.8	0.076
Colour: red	0.005	0.003	1.6	0.102
a1	0.125	0.009	13.5	<0.001
gender: boy	0.045	0.016	2.9	0.004
Run	-0.011	0.005	-2.4	0.018
Star_r	0.023	0.005	4.7	<0.001
OLB_I	0.005	0.003	4.7	<0.001

Table 9. 16: Random effects of m5

Groups	Name	Var	SD	Corr			
Child	Grand mean	0.148	0.022				
	a1	0.07	0.005	0.361			
	Congruence: con	0.016	0.0	0.066	0.068		
	Side: right	0.012	0.0	-0.09	0.22	-0.056	
	Colour: red	0.024	0.001	-0.43	0.113	-0.431	-0.617
Residual		0.24	0.058				

9.1.2.3 Means and standard deviations of model variabels

Table 9. 17: Mean and standard deviations of variables used in m5 and m8

Row	Condition	t0	t1	t2	t3	t6	t8
1	number of reaction times [n]	4388	3940	3724	3611	2586	1967
2	age [years]	9.2±0.5	9.4±0.5	9.6±0.5	10.1±0.6	10.7±0.5	11.6±0.5
3	gender [girls/boys]	35/41	29/38	28/37	28/33	18/26	15/19
4	20 m sprint [m/s]	4.1±0.4	4.1±0.4	4.1±0.4	4.1±0.4	4.2±0.4	4.3±0.4
5	star run [m/s]	1.8±0.2	1.8±0.2	1.8±0.2	1.9±0.2	2±0.2	2±0.3
6	ball push test [m]	3.6±0.7	3.7±0.7	3.7±0.6	3.8±0.7	4.1±0.7	4.8±0.8
7	one leg balance [log(s)]	2±0.8	2.2±0.7	2.1±0.8	2.1±0.8	2.4±0.6	2.5±0.7
8	6 min run [m]	859±125	879±135	846±139	868±136	864±146	844±177
9	standing long jump [cm]	108±19	108±20	110±20	112±21	114±21	124±21
4	blue left (incongruent)	1318±455	1206±426	1177±510	1094±442	1088±478	922±307
5	blue right (congruent)	1265±573	1202±580	1095±355	1051±434	1013±494	861±319
6	red left (congruent)	1231±505	1170±516	1095±524	1016±395	1049±499	902±392
7	red right (incongruent)	1303±481	1274±627	1210±693	1083±449	1119±524	900±293