UNIVERSITY OF POTSDAM

Faculty of Human Sciences

Research Focus Cognition Sciences

Division of Training and Movement Sciences

Secular trends, age, sex, and timing of school enrollment effects on physical fitness in children and adolescents

An academic thesis submitted to

the Faculty of Human Sciences of the University of Potsdam

for the degree

Doctor of Philosophy (Dr. phil.)

in the field of sports science

by

Thea Heidi Fühner

born in Rheine, Germany

Cologne, 11.07.2022

Principle supervisor:
Prof. Dr. Urs Granacher

Second supervisor:
Prof. Dr. Reinhold Kliegl

First reviewer:
Prof. Dr. Urs Granacher

Second reviewer:
Prof. Dr. Dirk Büsch

Date of disputation:
Potsdam, 20.03.2023

Published online on the
Publication Server of the University of Potsdam:
https://doi.org/10.25932/publishup-58864
https://nbn-resolving.org/urn:nbn:de:kobv:517-opus4-588643

Affidavits according to doctoral degree regulations (§ 4 (2), sentences No. 4 and 7) of the Faculty of Human Sciences, University of Potsdam:

Hereby, I declare that this thesis entitled 'Secular trends, age, sex, and timing of school enrollment effects on physical fitness in children and adolescents', or parts of the thesis have not yet been submitted for a doctoral degree to this or any other institution neither in identical nor in similar form. The work presented in this thesis is the original work of the author. I did not receive any help or support from commercial consultants. All parts or single sentences, which have been taken analogously or literally from other sources, are identified as citations. Additionally, significant contributions from co-authors to the article of this cumulative dissertation are acknowledged in the authors' contribution section.

______________________________ ______________________________

Place, Date **Thea Heidi Fühner**

Table of contents

Abstract

Background and objectives: The relevance of physical fitness for children's and adolescents' health is indisputable and it is crucial to regularly assess and evaluate children's and adolescents' individual physical fitness status to detect potential negative health consequences in time. Physical fitness tests are easy-to-administer, reliable, and valid which is why they should be widely used to provide information on performance development and health status of children and adolescents. When talking about development of physical fitness, two perspectives can be distinguished. One perspective is how the physical fitness status of children and adolescents changed / developed over the past decades (i.e., secular trends). The other perspective covers the analyses how physical fitness develops with increasing age due to growth and maturation processes. Although, the development of children's and adolescents' physical fitness has been extensively described and analyzed in the literature, still some questions remain to be uncovered that will be addressed in the present doctoral thesis.

Previous systematic reviews and meta-analyses have examined secular trends in children's and adolescents' physical fitness. However, considering that those analyses are by now 15 years old and that updates are available only to limited components of physical fitness, it is time to re-analyze the literature and examine secular trends for selected components of physical fitness (i.e., cardiorespiratory endurance, muscle strength, proxies of muscle power, and speed). Furthermore, the available studies on children's development of physical fitness as well as the effects of moderating variables such as age and sex have been investigated within a *long-term ontogenetic* perspective. However, the effects of age and sex in the transition from pre-puberty to puberty in the ninth year of life using a *short-term ontogenetic* perspective and the effect of timing of school enrollment on children's development of physical fitness have not been clearly identified. Therefore, the present doctoral thesis seeks to complement the knowledge of children's and adolescents' physical fitness development by updating secular trend analysis in selected components of physical fitness, by examining *short-term ontogenetic* cross-sectional developmental differences in children`s physical fitness, and by comparing physical fitness of older- and younger-than-keyage children versus keyage-children. These findings provide valuable information about children's and adolescents' physical fitness development to help prevent potential deficits in physical fitness as early as possible and consequently ensure a holistic development and a lifelong healthy life.

Methods: Initially, a systematic review to provide an 'update' on secular trends in selected components of physical fitness (i.e., cardiorespiratory endurance, relative muscle strength, proxies of muscle power, speed) in children and adolescents aged 6 to 18 years was conducted using the Preferred Reporting Items for Systematic Reviews and Meta-Analysis statement guidelines. To examine *short-term ontogenetic* cross-sectional developmental differences and to compare physical fitness of older- and younger-than-keyage children versus keyage-children physical fitness data of 108,295 keyage-children (i.e., aged 8.00 to 8.99 years), 2,586 younger-than-keyage children (i.e., aged 7.00 to 7.99 years), and 26,540 older-than-keyage children (i.e., aged 9.00 to 9.99 years) from the third grade were analyzed. Physical fitness was assessed through the EMOTIKON test battery measuring cardiorespiratory endurance (i.e., 6-min-run test), coordination (i.e., star-run test), speed (i.e., 20-m linear sprint test), and proxies of lower (i.e., standing long jump test) and upper limbs (i.e., ball-push test) muscle power. Statistical inference was based on Linear Mixed Models.

Results: Findings from the systematic review revealed a large initial improvement and an equally large subsequent decline between 1986 and 2010 as well as a stabilization between 2010 and 2015 in cardiorespiratory endurance, a general trend towards a small improvement in relative muscle strength from 1972 to 2015, an overall small negative quadratic trend for proxies of muscle power from 1972 to 2015, and a small-to-medium improvement in speed from 2002 to 2015. Findings from the cross-sectional studies showed that even in a single prepubertal year of life (i.e., ninth year) physical fitness performance develops linearly with increasing chronological age, boys showed better performances than girls in all physical fitness components, and the components varied in the size of sex and age effects. Furthermore, findings revealed that older-than-keyage children showed poorer performance in physical fitness compared to keyage-children, older-than-keyage girls showed better performances than older-than-keyage boys, and younger-than-keyage children outperformed keyage-children.

Conclusions: Due to the varying secular trends in physical fitness, it is recommended to promote initiatives for physical activity and physical fitness for children and adolescents to prevent adverse effects on health and well-being. More precisely, public health initiatives should specifically consider exercising cardiorespiratory endurance and muscle strength because both components showed strong positive associations with markers of health. Furthermore, the findings implied that physical education teachers, coaches, or researchers can utilize a proportional adjustment to individually interpret physical fitness of prepubertal school-aged children. Special attention should be given to the promotion of physical fitness of older-than-keyage children because they showed poorer performance in physical fitness than keyage-children. Therefore, it is necessary to specifically consider this group and provide additional health and fitness programs to reduce their deficits in physical fitness experienced during prior years to guarantee a holistic development.

Keywords: physical fitness, children, adolescents, primary school, secular trends, age, sex, timing of school enrollment

Zusammenfassung

Hintergrund und Ziele: Die Relevanz der körperlichen Fitness für die Gesundheit von Kindern und Jugendlichen ist unbestritten und es ist von entscheidender Bedeutung, den individuellen körperlichen Fitnesszustand von Kindern und Jugendlichen regelmäßig zu untersuchen und zu bewerten, um mögliche negative gesundheitliche Folgen rechtzeitig zu erkennen. Tests zur Bewertung der körperlichen Fitness sind einfach durchzuführen, reliabel und valide, weshalb sie in großem Umfang eingesetzt werden sollten, um Informationen über die Entwicklung der körperlichen Fitness und den Gesundheitszustand von Kindern und Jugendlichen zu erhalten. Bei der Betrachtung der Entwicklung der körperlichen Fitness lassen sich zwei Perspektiven unterscheiden. Die eine Perspektive untersucht, wie sich die körperliche Fitness von Kindern und Jugendlichen in den letzten Jahrzehnten verändert bzw. entwickelt hat (sog. säkulare Trends). Die andere Perspektive analysiert, wie sich die körperliche Fitness mit zunehmendem Alter aufgrund von Wachstums- und Reifungsprozessen entwickelt. Obwohl die Entwicklung der körperlichen Fitness von Kindern und Jugendlichen in der Literatur ausführlich beschrieben und analysiert wurde, sind noch einige Fragen offen, welche in der vorliegenden Doktorarbeit adressiert werden.

Frühere systematische Überblicksbeiträge und Meta-Analysen haben säkulare Trends in der körperlichen Fitness von Kindern und Jugendlichen untersucht. Da diese Analysen jedoch mittlerweile 15 Jahre alt sind und Aktualisierungen nur für ausgewählte Komponenten der körperlichen Fitness zur Verfügung stehen, ist es notwendig, die Literatur neu zu analysieren und säkulare Trends für ausgewählte Komponenten der körperlichen Fitness (d. h. aerobe Ausdauer, relative Muskelkraft, Schnellkraft und Schnelligkeit) zu untersuchen. Darüber hinaus wurden die verfügbaren Studien über die Entwicklung der körperlichen Fitness von Kindern sowie die Effekte moderierender Variablen wie Alter und Geschlecht in einer langfristigen ontogenetischen Perspektive untersucht. Die Effekte von Alter und Geschlecht beim Übergang von der Vorpubertät zur Pubertät im neunten Lebensjahr unter einer kurzfristigen ontogenetischen Perspektive und die Effekte des Zeitpunkts der Einschulung auf die Entwicklung der körperlichen Fitness von Kindern sind jedoch bislang nicht eindeutig identifiziert worden.

Die vorliegende Doktorarbeit hat daher das Ziel, den Wissensstand über die Entwicklung der körperlichen Fitness von Kindern und Jugendlichen zu ergänzen, indem säkulare Trendanalysen ausgewählter Komponenten der körperlichen Fitness aktualisiert werden, kurzfristige ontogenetische Entwicklungsunterschiede in der körperlichen Fitness von Kindern im Querschnitt untersucht werden und die körperliche Fitness von älteren und jüngeren Kindern im Vergleich zu Stichtagskindern analysiert wird. Diese Erkenntnisse liefern wertvolle Informationen über die Entwicklung der körperlichen Fitness von Kindern und Jugendlichen, um möglichen Defiziten in der körperlichen Fitness so früh wie möglich vorzubeugen und damit eine ganzheitliche Entwicklung und ein lebenslanges gesundes Leben zu gewährleisten.

Methodik: Zunächst wurde ein systematischer Überblicksbeitrag verfasst, um den Wissenstand der säkularen Trends bei ausgewählten Komponenten der körperlichen Fitness (d. h. aerobe Ausdauer, relative Muskelkraft, Schnellkraft, Schnelligkeit) bei Kindern und Jugendlichen gemäß den PRISMA Richtlinien zu aktualisieren. Um kurzfristige ontogenetische Entwicklungsunterschiede im Querschnitt zu untersuchen und die körperliche Fitness von älteren und jüngeren Kindern im Vergleich zu Stichtagskindern zu analysieren, wurden die Daten zur körperlichen Fitness von 108.295 Stichtagskindern (d. h. im Alter von 8,00 bis 8,99 Jahren), 2.586 jüngeren Kindern (d. h. im Alter von 7,00 bis 7,99 Jahren) und 26.540 älteren Kindern (d. h. im Alter von 9,00 bis 9,99 Jahren) aus der dritten Klasse analysiert. Die körperliche Fitness wurde anhand der EMOTIKON-Testbatterie zur Messung der aeroben Ausdauer (d. h. 6-min-Lauf), der Koordination (d. h. Sternlauf), der Schnelligkeit (d. h. 20-m-Sprint) sowie der Schnellkraft der unteren (d. h. Standweitsprung) und oberen Extremitäten (d. h. Medizinballstoßen) erhoben. Die statistische Inferenz basierte auf Linear Mixed Models.

Ergebnisse: Die Ergebnisse des systematischen Überblickbeitrags zeigten eine starke anfängliche Verbesserung und einen ebenso starken Rückgang zwischen 1986 und 2010 sowie eine Stabilisierung zwischen 2010 und 2015 bei der aeroben Ausdauer. Zudem war ein allgemeiner Trend zu einer geringen Verbesserung der relativen Muskelkraft zwischen 1972 und 2015, ein insgesamt geringer negativer quadratischer Trend bei der Schnellkraft zwischen 1972 und 2015 sowie eine geringe bis mittlere Verbesserung der Schnelligkeit zwischen 2002 und 2015 zu verzeichnen. Die Ergebnisse der Querschnittsstudien legten dar, dass sich die körperliche Fitness selbst in einem einzigen präpubertären Lebensjahr (d. h. neuntes Lebensjahr) mit zunehmendem chronologischen Alter linear entwickelt, dass Jungen in allen untersuchten Komponenten der körperlichen Fitness bessere Leistungen zeigten als Mädchen und dass sich die Komponenten der körperlichen Fitness in der Stärke der Alters- und Geschlechtseffekte stark unterschieden. Darüber hinaus zeigten die Ergebnisse, dass Kinder, die älter als Stichtagskinder sind, im Vergleich zu Stichtagskindern schlechtere Leistungen zeigten; Mädchen, die älter als Stichtagskinder sind, bessere Leistungen zeigten als Jungen, die älter als Stichtagskinder sind; und Kinder, die jünger als Stichtagskinder sind, bessere Leistungen zeigten als Stichtagskinder.

Schlussfolgerung: Aufgrund der unterschiedlichen säkularen Trends bei der körperlichen Fitness wird empfohlen, Initiativen für körperliche Aktivität und körperliche Fitness bei Kindern und Jugendlichen zu fördern, um negative Auswirkungen auf die Gesundheit und das Wohlbefinden zu verhindern. Insbesondere sollten Initiativen im Bereich der öffentlichen Gesundheit speziell das Training der aeroben Ausdauer und der Muskelkraft berücksichtigen, da beide Komponenten starke positive Assoziationen mit Gesundheitsmarkern aufweisen. Darüber hinaus legen die Ergebnisse nahe, dass Sportlehrkräfte, Trainer und Trainerinnen oder Wissenschaftler und Wissenschaftlerinnen eine proportionale Anpassung nutzen können, um die körperliche Fitness von Kindern im vorpubertären Schulalter individuell zu interpretieren. Ein besonderes Augenmerk sollte auf die Förderung der körperlichen Fitness von Kindern, die älter als Stichtagskinder sind, gelegt werden, da diese schlechtere Leistungen in der körperlichen Fitness zeigten als Stichtagskinder. Daher ist es notwendig, diese Gruppe besonders zu berücksichtigen und zusätzliche Gesundheits- und Fitnessprogramme anzubieten, um ihre Defizite in der körperlichen Fitness abzubauen und eine ganzheitliche Entwicklung zu gewährleisten.

Schlüsselwörter: körperliche Fitness, Kinder, Jugendliche, Grundschule, säkulare Trends, Alter, Geschlecht, Zeitpunkt der Einschulung

List of figures

List of publications

The present cumulative thesis is based on the following publications.

Publication I Führer, T., Kliegl, R., Arntz, F., Kriemler, S., & Granacher, U. (2021a). An update on secular trends in physical fitness of children and adolescents from 1972 to 2015: a systematic review. *Sports Medicine, 51*, 303-320. https://doi.org/10.1007/s40279-020-01373-x

The final publication is available at:

https://link.springer.com/article/10.1007/s40279-020-01373-x

Publication II Führer, T., Granacher, U., Golle, K., & Kliegl, R. (2021b). Age and sex effects in physical fitness components of 108,295 third graders including 515 primary schools and 9 cohorts. *Scientific Reports, 11*, 17566. https://doi.org/10.1038/s41598-021-97000-4

The final publication is available at:

https://www.nature.com/articles/s41598-021-97000-4

Publication III Führer, T., Granacher, U., Golle, K., & Kliegl, R. (2022). Effect of timing of school enrollment on physical fitness in third graders. *Scientific Reports, 12,* 7801. https://doi.org/10.1038/s41598-022-11710-x

The final publication is available at:

https://www.nature.com/articles/s41598-022-11710-x

Abbreviations

BMI	Body Mass Index
CPs	correlation parameters
KiGGS	German Health Interview and Examination Survey for Children and Adolescents
LMM	Linear Mixed Model
OTK	older-than-keyage children
p.a.	per annum
powerLOW	proxies of lower limbs muscle power
powerUP	proxies of upper limbs muscle power
PRISMA	Preferred Reporting Items for Systematic Reviews and Meta-Analysis
PWC	Physical Work Capacity
VCs	variance components
$\dot{V}O_{2peak}$	peak oxygen uptake at the end of the load without flattening of oxygen uptake
WHO	World Health Organization
YTK	younger-than-keyage children

1. General introduction

Childhood is an important developmental period in which fundamental movement skills are obtained through daily physical activity to acquire motor skill competence and movement confidence (Faigenbaum & Myer, 2012). Therefore, according to evidence-based research, the World Health Organization (WHO) postulates that children and adolescents aged 5 to 17 years should be physically active at least an average of[1] 60 minutes a day at moderate-to-vigorous intensity (Bull et al., 2020; World Health Organization, 2010). The WHO states that 60 minutes of daily physical activity are enough to provide significant health benefits (e.g., physical fitness, bone health, cognitive outcomes [i.e., academic performance, executive function], sleep) and to reduce health risks (e.g., adiposity, mental symptoms of anxiety and depression, cardiometabolic health [i.e., blood pressure, insulin resistance]) (Bull et al., 2020). However, recently published studies showed that only a small percentage of children and adolescents (~ 20%) worldwide met the recommended level of 60 min physical activity per day (Guthold et al., 2020; Hallal et al., 2012; Kalman et al., 2015). As childhood is an important developmental period in which fundamental movement skills are obtained through daily physical activity to acquire motor skill competence and movement confidence, low levels of physical activity prevent the acquirement of these skills (Faigenbaum & Myer, 2012). Consequently, children and adolescents who do not acquire these competencies due to physical inactivity are at higher risk of having disease risk factors and developing negative health outcomes later in life (Bull et al., 2020; Faigenbaum & Myer, 2012). Therefore, high amounts of physical activity are an important foundation for lifelong health.

Findings from cross-sectional (Katzmarzyk et al., 1998; Mayorga-Vega et al., 2019; Wrotniak et al., 2006) and longitudinal studies (Larsen et al., 2015) revealed that physical activity is strongly related to physical fitness. According to Caspersen et al. (1985), physical fitness can be classified as health- (i.e., cardiorespiratory endurance, muscular endurance, muscular strength, body composition, and flexibility) and skill-related (i.e., agility, balance, coordination, speed, [muscle] power, and reaction time) components of physical fitness. Foremost cardiorespiratory endurance (Mintjens et al., 2018) and muscle strength (García-Hermoso et al., 2019; Rodrigues de Lima et al., 2020; Smith et al., 2014) showed the strongest associations with indicators of health in children and adolescents. Accordingly, the relevance of physical fitness for children's and adolescents' health is indisputable and it is crucial to regularly assess

[1] In the updated version of the 2010 *Global Recommendations on Physical Activity for Health* (World Health Organization, 2010) the WHO changed the wording 'at least' to 'at least an average of' (Bull et al., 2020). Therefore, the new wording of 'at least an average of' is used in this present doctoral thesis.

and evaluate children's and adolescents' individual physical fitness status to detect potential negative health consequences in time. Physical fitness tests are easy-to-administer, reliable, and valid which is why they should be widely used to provide information on performance development and health status of children and adolescents (Pate, 1991).

Although the development of children's and adolescents' physical fitness has been extensively described and analyzed in literature, still some questions remain unknown that will be addressed in the present doctoral thesis to provide additional information about children's and adolescents' physical fitness development. Those findings can be used to prevent potential deficits in physical fitness as early as possible and consequently ensure a holistic development and lifelong health.

The first question addresses secular trends. Secular trend analyses examine the extent to which physical fitness changes over time (Albrecht et al., 2016). Probably the best-known studies about secular trends in physical fitness are those conducted by Tomkinson (2007) and Tomkinson and Olds (2007). They summarized the existing literature on secular trends for cardiorespiratory endurance, proxies of muscle power, and speed for the timespan between 1958 and 2003. Their analysis considered physical fitness data of 25,000,000 to 50,000,000 children and adolescents aged 6 to 19 years from all over the world. Between 1958 and 2003, cardiorespiratory endurance declined by - 0.36% per annum (p.a.) (Tomkinson & Olds, 2007) whereas proxies of muscle power and speed increased by + 0.03% p.a. and + 0.04% p.a., respectively (Tomkinson, 2007). Recently these two well-known studies were updated for cardiorespiratory endurance and absolute muscle strength by analyzing data on the 20-m shuttle run test (Tomkinson et al., 2019) and handgrip strength (Dooley et al., 2020). The updated analyses showed that the international rate of decline in cardiorespiratory endurance has diminished and stabilized since 2000 (Tomkinson et al., 2019) and the international rate of improvement in handgrip strength progressively increased between 1967 and 2017 (Dooley et al., 2020).

Particularly the two popular studies conducted by Tomkinson (2007) and Tomkinson and Olds (2007) were the starting point for among others international promotion of physical activity and physical fitness in children and adolescents, especially by the WHO (2018). Considering the extensive promotion of physical activity and physical fitness by the WHO in recent years, it seems necessary to conduct regular updates of these secular trend analyses to evaluate if / how secular trends might have changed in the last years. For example, in case that children's and adolescents' physical fitness performance decreased over time, this alarming result could be used to design national or international initiatives in physical fitness promotion to ensure a holistic development and a lifelong healthy life. Since these updates have only been conducted for cardiorespiratory endurance using estimated $\dot{V}O_{2peak}$ from the 20-m shuttle run test (Tomkinson et al., 2019) and absolute muscle strength data from handgrip dynamometry (Dooley et al., 2020), secular trends for other components of physical fitness such as relative muscle strength, proxies of muscle power, or speed remain unknown. Furthermore, it is also necessary to widen the perspective on cardiorespiratory endurance, i.e., to include different tests to the 20-m shuttle run test in secular trend analyses. Consequently, there is a need to fill this gap in the literature and complement the body of knowledge regarding secular trends in selected components of physical fitness.

As already mentioned, secular trend analyses examine the extent to which physical fitness changes / develops[2] over time. Thus, when talking about secular trends in physical fitness the perspective is on how physical fitness status of children and adolescents developed over the past decades (i.e., increase, stagnation, or decrease). However, physical fitness cannot only change / develop over time but also develops[2] with increasing age particularly during childhood and adolescence due to growth (i.e., increasing body mass and body height) and maturation processes (Malina et al., 2004). These growth and maturation induced developmental gains in physical fitness are foremost modulated by age and sex. Those two moderators will be addressed in the second question of this doctoral thesis.

There is evidence that physical fitness develops progressively during childhood and adolescence and is modulated by age and sex (De Miguel-Etayo et al., 2014; Golle et al., 2015; Malina et al., 2004; Niessner et al., 2020; Santos et al., 2014; Tambalis et al., 2016; Tomkinson et al., 2018). Accordingly, many studies worldwide (Castro-Piñero et al., 2010; Catley & Tomkinson, 2013; De Miguel-Etayo et al., 2014; Dobosz et al., 2015; Golle et al., 2015; Niessner et al., 2020; Ortega et al., 2011; Roriz de Oliveira et al., 2014; Santos et al., 2014; Tambalis et al., 2016; Tomkinson et al., 2018) revealed developmental gains in physical fitness from childhood to adolescence (De Miguel-Etayo et al., 2014; Dobosz et al., 2015; Golle et al., 2015; Niessner et al., 2020; Ortega et al., 2011; Santos et al., 2014; Tambalis et al., 2016). Boys show better performances compared to girls during childhood and adolescence in most components of physical fitness (Castro-Piñero et al., 2010; Catley & Tomkinson, 2013; De Miguel-Etayo et al., 2014; Dobosz et al., 2015; Fraser et al., 2019; Lisowski et al., 2020; Ortega

[2] Normally, the word *development* relates to a longitudinal process and intraindividual growth. However, in this present doctoral thesis *development* refers on the one hand to secular trends (i.e., development of physical fitness over time) and on the other hand describes the ontogenetic cross-sectional development of physical fitness with increasing age (i.e., within-year developmental profiles).

et al., 2011; Roriz de Oliveira et al., 2014; Santos et al., 2014; Tambalis et al., 2016; Tomkinson et al., 2018), except for balance (De Miguel-Etayo et al., 2014; Dobosz et al., 2015) and flexibility (Catley & Tomkinson, 2013; De Miguel-Etayo et al., 2014; Dobosz et al., 2015; Golle et al., 2015; Lisowski et al., 2020; Ortega et al., 2011; Roriz de Oliveira et al., 2014; Santos et al., 2014; Tambalis et al., 2016; Tomkinson et al., 2018).

Those differences between boys and girls are especially evident during puberty (Laurson et al., 2022) and puberty-related differences between boys and girls are well documented within the literature. All of the above studies have in common that they examined a wide age range spanning childhood and adolescence (e.g., 9 to 17 years (Tomkinson et al., 2018)), that is a *long-term ontogenetic* perspective. However, a *short-term ontogenetic* perspective (i.e., focusing on a single year of life) might unveil different growth trajectories that might be overlooked within a *long-term ontogenetic* perspective. For example, the situation of boys' and girls' divergence during puberty might be different during prepuberty. Girls visibly mature about two years earlier than boys (Stratton & Oliver, 2020), but sex hormones rise already considerably starting at eight years of age (Marshall & Tanner, 1969, 1970; Rosenfield et al., 2009). If the early rise of sex hormones is related to body composition, the question arises whether there is evidence for faster development of girls compared to boys in this single year of life (i.e., ninth year), leading to a convergence of performance in physical fitness during the transition from pre-puberty to puberty. A *long-term ontogenetic* perspective on this question might be too 'long' to unveil any differences, and a *short-term ontogenetic* perspective might be the better approach. However, the effects of age and sex in physical fitness components in the transition from pre-puberty to puberty in the ninth year of life (i.e., developmental differences) have not yet been investigated within a *short-term ontogenetic* perspective. Thus, the second question of this doctoral thesis will deal with this *short-term ontogenetic* perspective. These findings can further the understanding of children's development of physical fitness and can be helpful for an individual evaluation of children's physical fitness status.

As mentioned above, children's physical fitness develops continuously during childhood and adolescence. Particularly, schools with their physical education play an important role in children´s acquirement of fundamental movement skills and obtainment of motor skill competence and movement confidence (Faigenbaum & Myer, 2012) because schools reach all children during their development, growth, and formation of habits. Within schools, children are matched into grades based on their date of birth and a country-specific school enrollment date. For instance, in the Federal State of Brandenburg, Germany September 30th is the official date for school enrollment. Accordingly, children who are between 6.00 and 6.99 years old on

September 30th are enrolled to school (i.e., keyage-children). However, due to individual developmental processes, parents have the option to enroll a gifted child earlier or a slow developing child later (Ministerium für Bildung Jugend und Sport, 2018). Moreover, gifted children can skip a school year or children with learning deficits may have to repeat a school year. All these aspects imply that the age range within school grades can vary considerably including older-than-keyage (OTK) and younger-than-keyage (YTK) children. As already mentioned, physical fitness develops with increasing age during childhood and adolescents. However, the effect of timing of school enrollment on physical fitness development remains unclear and it has not yet been investigated how physical fitness of OTK and YTK children develops compared to keyage-children. Therefore, this will be addressed in the third question of the present doctoral thesis. These findings can be interesting for physical education teachers to individually evaluate student's physical fitness considering their age, sex, and timing of school enrollment. Based on the results additional health and fitness programs can be designed e.g., to compensate potential deficits in physical fitness as early as possible and consequently ensure a holistic development and a lifelong healthy life.

Taken together, previous systematic reviews and meta-analyses have examined secular trends in children's and adolescents' physical fitness (Dooley et al., 2020; Tomkinson & Olds, 2007; Tomkinson, 2007; Tomkinson et al., 2019). However, considering that those analyses are by now 15 years old and that updates are available only to limited components of physical fitness, it is time to re-analyze the literature and examine secular trends for selected components of physical fitness (i.e., cardiorespiratory endurance, muscle strength, proxies of muscle power, and speed). Furthermore, the available studies on children's development of physical fitness as well as the effects of moderating variables such as age and sex (De Miguel-Etayo et al., 2014; Golle et al., 2015; Malina et al., 2004; Niessner et al., 2020; Santos et al., 2014; Tambalis et al., 2016; Tomkinson et al., 2018) have been investigated within a *long-term ontogenetic* perspective. However, the effects of age and sex in the transition from pre-puberty to puberty in the ninth year of life using a *short-term ontogenetic* perspective and the effect of timing of school enrollment on children's development of physical fitness have not been clearly identified. Therefore, the present cumulative thesis comprises of two parts. The first part consists of a systematic review on an update on secular trends in selected components of physical fitness (i.e., cardiorespiratory endurance, relative muscle strength, proxies of muscle power, speed) in children and adolescents aged 6 to 18 years (Fühner et al., 2021a). The second part involves two cross-sectional studies. The first addresses *short-term ontogenetic cross-sectional* developmental differences in physical fitness for five tests tapping health- and skill-related components of physical fitness in a large sample of 108,295 German eight-year old children (Fühner et al., 2021b). The second cross-sectional study compares physical fitness of OTK and YTK children versus keyage-children in a sample of German primary school children (Fühner et al., 2022) to further understand children's development of physical fitness. These findings can provide valuable information about children's and adolescents' physical fitness development to help prevent potential deficits in physical fitness as early as possible and consequently ensure a holistic development and a lifelong healthy life.

2. Literature Review

This literature review provides information on physical activity status of children and adolescents nationally and internationally and outlines the association between levels of physical activity and physical fitness as well as the relevance of physical fitness. In addition, it summarizes the findings on previous research regarding secular trends in physical fitness and effects of age, sex, and timing of school enrollment on children`s development of physical fitness.

2.1 Physical activity in children and adolescents

Physical activity refers to "any bodily movement produced by skeletal muscles that results in energy expenditure" (Caspersen et al., 1985, p. 129). According to evidence-based research, the WHO postulates that children and adolescents aged 5 to 17 years should be physically active for at least an average of 60 minutes a day at moderate-to-vigorous intensity (Bull et al., 2020; World Health Organization, 2010). The WHO states that 60 minutes of daily physical activity are enough to provide significant health benefits (e.g., physical fitness, bone health, cognitive outcomes [i.e., academic performance, executive function], sleep) and to reduce health risks (e.g., adiposity, mental symptoms of anxiety and depression, cardiometabolic health [i.e., blood pressure, insulin resistance]) (Bull et al., 2020). In addition, resistance training three times a week is recommended to strengthen muscles and bones (Bull et al., 2020; World Health Organization, 2010). The national physical activity recommendations for Germany of the Robert Koch Institute for children and adolescents aged 6 to 18 years even refer to a daily physical activity time of 90 minutes. In addition, resistance training to strengthen the large muscle groups is recommended two to three days per week (Rütten & Pfeifer, 2016).

The second wave (2014 to 2017) of the *German Health Interview and Examination Survey for Children and Adolescents* (KiGGS) showed that only 22.4% of German girls and 29.4% of German boys aged 3 to 17 years met the WHO guidelines. While the percentage of children who met the recommended level of at least an average of 60 min physical activity per day was comparatively high in children aged 3 to 6 years (♂ = 48.9%; ♀ = 42.5%), it decreased with increasing age (♂ = 16.0%; ♀ = 7.5% in adolescents aged 14 to 17 years). In addition, comparison of physical activity data from KiGGS wave 1 (2009 to 2012) with that from the second wave (2014 to 2017) showed that the percentage of children and adolescents who met the recommended level of at least an average of 60 min physical activity per day was stable among boys (29.7% to 29.4%), whereas it decreased among girls (25.9% to 22.4%) (Finger et al., 2018). The low percentage of children and adolescents who meet the WHO recommendations is not only observed in Germany but also internationally. Recently published studies investigated the percentage of children and adolescents worldwide who met the WHO recommendations and showed that also internationally only a small percentage of children and adolescents (~ 20%) met the recommended level of 60 min physical activity per day (Guthold et al., 2020; Hallal et al., 2012; Kalman et al., 2015). In addition, only a small percentage of European adolescents aged 15 to 17 (♂ = 24.7%; ♀ = 13.8%) met the recommended level of muscle-strengthening activity three times a week (Bennie et al., 2021).

In the case of non-compliance with the WHO recommendations (World Health Organization, 2010), the term ‘exercise deficit disorder’ has been established in the Anglo-American language to describe all negative health consequences of low levels of physical activity (Faigenbaum & Myer, 2012). As childhood is an important developmental period in which fundamental movement skills are obtained through daily physical activity to acquire motor skill competence and movement confidence, low levels of physical activity prevent the acquirement of these skills. Consequently, children and adolescents who do not acquire these competencies due to physical inactivity are at higher risk of having disease risk factors and developing negative health outcomes later in life (Bull et al., 2020; Faigenbaum & Myer, 2012). The lack of physical activity has far-reaching consequences for children and adolescents, as physical activity behavior during the formative stages of development (i.e., childhood and adolescence) is relatively stable and can affect behavior as an adult (i.e., a physically active / inactive child is also very likely to be a physically active / inactive adult) (Fraser et al., 2017, 2021; García-Hermoso et al., 2019; Ortega et al., 2008; Poitras et al., 2016; Telama et al., 2014). For instance, based on a 27-year longitudinal study with 3,596 Finnish boys and girls aged 3 to 18 years, Telama et al. (2014) reported that physical activity levels during childhood and adolescence extended into adulthood with moderate-to-high stability. Fraser et al. (2017) pointed to a similar finding in a longitudinal study of muscle strength with 623 boys and girls at ages 9, 12, and 15 years. A follow-up measurement after 20 years showed that the level of muscle strength in childhood extended into adulthood with moderate stability. Thus, high amounts of physical activity are an important foundation for a lifelong healthy life.

There is evidence from cross-sectional (Katzmarzyk et al., 1998; Mayorga-Vega et al., 2019; Wrotniak et al., 2006) and longitudinal studies (Larsen et al., 2015) of positive correlations between levels of physical activity and physical fitness. For instance, Wrotniak et al. (2006) reported that levels of moderate-to-vigorous physical activity (measured by accelerometers) showed positive correlations with various components of physical fitness (muscular strength [i.e., standing long jump] r = 0.40; speed [i.e., running speed] r = - 0.36) in children aged 8 to 10 years. These results were proved in a 3-year longitudinal study with children aged 6 to 12 years that showed positive associations between levels in moderate-to-vigorous physical activity (measured by accelerometers) and different components of physical fitness (muscle strength [i.e., handgrip] β = 0.06); proxies of muscle power [i.e., vertical jump] β = 0.04) (Larsen et al., 2015).

2.2 Relevance of children’s and adolescents’ physical fitness

Evidence-based research shows that physical fitness is an important indicator of children’s and adolescents’ health (Ortega et al., 2008). According to Caspersen et al. (1985), physical fitness can be classified as health- (i.e., cardiorespiratory endurance, muscular endurance, muscular strength, body composition, and flexibility) and skill-related (i.e., agility, balance, coordination, speed, [muscle] power, and reaction time) components of physical fitness. Foremost cardiorespiratory endurance (Mintjens et al., 2018) and muscle strength (García-Hermoso et al., 2019; Rodrigues de Lima et al., 2020; Smith et al., 2014) showed the strongest associations with indicators of health in children and adolescents. For instance, in a systematic review, Mintjens et al. (2018) showed that performance levels in cardiorespiratory endurance were positively associated with body mass index (BMI), waist circumference, body fatness, and prevalence of metabolic syndrome. In addition, García-Hermoso et al. (2019) showed that measures of muscle strength were positively related with BMI, skinfold thickness, insulin resistance, triglycerides, cardiovascular disease risk score, and bone mineral density.

Accordingly, the relevance of physical fitness for children’s and adolescents’ health is indisputable and it is crucial to regularly assess and evaluate children’s and adolescents’ individual physical fitness status to detect potential negative health consequences in time. Physical fitness tests are easy-to-administer, reliable, and valid which is why they should be widely used to provide information on performance development and health status of children and adolescents (Pate, 1991).

Normally, the word *development* relates to a longitudinal process and intraindividual growth. However, in this present doctoral thesis *development* refers on the one hand to secular trends (i.e., developmental changes of physical fitness over time) and on the other hand describes the *ontogenetic cross-sectional* development of physical fitness with increasing age (i.e., within-year developmental profiles). Thus, when talking about secular trends in physical fitness the perspective is on how physical fitness status of children and adolescents developed over the past decades (i.e., increase, stagnation, or decrease). In addition, physical fitness cannot only change / develop over time but also develops with increasing age particularly during childhood and adolescence due to growth (i.e., increasing body mass and body height) and maturation processes (Malina et al., 2004).

The following chapters briefly summarize previous research with regard to secular trends in physical fitness and the effects of age, sex, and timing of school enrollment on children`s development of physical fitness.

2.2.1 Secular trends in physical fitness

According to Faigenbaum et al. (2018), low levels of physical activity (i.e., exercise deficit disorder) negatively impact physical fitness. Thus, it is not surprising that there is already evidence in the literature of secular trends in physical fitness. Secular trend analyses examine the extent to which physical fitness changes over time (Albrecht et al., 2016).

Probably the best-known studies about secular trends in physical fitness are those conducted by Tomkinson (2007) and Tomkinson and Olds (2007). They summarized the existing literature on secular trends for cardiorespiratory endurance, proxies of muscle power, and speed for the timespan between 1958 and 2003. Their analysis considered physical fitness data of 25,000,000 to 50,000,000 children and adolescents aged 6 to 19 years from all over the world. Between 1958 and 2003, cardiorespiratory endurance declined by - 0.36% p.a. A detailed analysis revealed that between 1958 and 1970 cardiorespiratory endurance first increased by + 0.61% p.a. before declining by - 0.54% p.a. (Tomkinson & Olds, 2007). In contrast, proxies of muscle power and speed increased by + 0.03% p.a. and + 0.04% p.a. between 1958 and 2003 (Tomkinson, 2007). However, a detailed analysis again revealed different patterns over time showing that between 1958 and 1985 both components first increased (proxies of muscle power [+ 0.44% p.a.]; speed [+ 0.27% p.a.]) before both components declined (proxies of muscle power [- 0.20% p.a.]; speed [- 0.08% p.a.]) (Tomkinson, 2007).

Recently these two well-known studies were updated for cardiorespiratory endurance using data on the 20-m shuttle run test (Tomkinson et al., 2019), for muscle strength using data on handgrip strength (Dooley et al., 2020) and sit-up test (Kaster et al., 2020)[3], and for proxies of muscle

[3] The updates on secular trends in physical fitness conducted by Kaster et al. (Kaster et al., 2020) and Tomkinson et al. (Tomkinson et al., 2021) were not yet published when the first publication (Fühner et al., 2021a) was written.

power using data on standing long jump test (Tomkinson et al., 2021)[3]. In their update on cardiorespiratory endurance (i.e., 20-m shuttle run test), they summarized data for 965,264 children and adolescents aged 9 to 17 years living in 19 different countries between 1981 and 2014. Their results showed that the international rate of decline in cardiorespiratory endurance has diminished and stabilized since 2000 (Tomkinson et al., 2019). Furthermore, a recently published update of secular trends in muscle strength using handgrip strength data of 2,216,320 children and adolescents aged 9 to 17 years living in 19 countries showed that the international rate of improvement in handgrip strength continuously improved between 1967 and 2017 (Dooley et al., 2020). In addition, the analysis of secular trends on muscle strength (i.e., sit-up test) including data for 9,939,289 children and adolescents aged 9 to 17 years living in 31 countries showed that internationally rate of muscle strength improvement has slowed between 1964 to 2000, then stabilized until 2010, before declining (Kaster et al., 2020). Moreover, the update on secular trends in proxies of muscle power (i.e., standing long jump test) including data for 10,940,801 children and adolescents aged 9 to 17 years living in 29 countries showed that the international rate of proxies of muscle power improvement between 1960s and 1980s has slowed in 1990s, before declining (Tomkinson et al., 2021).

Particularly the two popular studies conducted by Tomkinson (2007) and Tomkinson and Olds (2007) were the starting point for e.g. international promotion of physical activity and physical fitness in children and adolescents, especially by the WHO (2018). Considering the extensive promotion of physical activity and physical fitness by the WHO in recent years, it seems necessary to conduct regular updates of these secular trend analyses. However, these updates have only been conducted for cardiorespiratory endurance using estimated $\dot{V}O_{2peak}$ from the 20-m shuttle run test (Tomkinson et al., 2019), absolute muscle strength data from handgrip dynamometry (Dooley et al., 2020), relative muscle strength data from sit-up test (Kaster et al., 2020), and proxies of muscle power data using standing long jump test (Tomkinson et al., 2021). Thus, secular trends for other components of physical fitness such as relative muscle strength or speed remain unknown. Furthermore, it is also necessary to widen the perspective on cardiorespiratory endurance and proxies of muscle power, i.e., to consider different tests to the 20-m shuttle run test or the standing long jump test in the analyses. Consequently, there is a need to fill this gap in the literature and complement the body of knowledge regarding secular trends in selected components of physical fitness.

Since the publication of the two popular studies by Tomkinson (2007) and Tomkinson and Olds (2007) numerous original studies were published analyzing secular trends in various components of physical fitness, over different time periods, in different countries around the world, and in diverse samples (for an overview of these studies see Publication I). For instance, Venckunas et al. (2017) examined secular trends in various components of physical fitness (i.e., balance, flexibility, muscular strength, proxies of muscle power, agility, and cardiorespiratory endurance) in 16,199 Lithuanian children and adolescents aged 11 to 18 years between 1992, 2002, and 2012. They reported decreases in cardiorespiratory endurance (i.e., 20-m shuttle run test) and proxies of lower limbs muscle power (i.e., standing long jump test) between 1992, 2002, and 2012 for all age groups and both sexes (Venckunas et al., 2017). In contrast, based on data from the KiGGS study, Albrecht et al. (2016) stated that cardiorespiratory endurance (i.e., PWC 170 cycle ergometer test) increased for 1,694 German adolescents aged 11 to 13 years for nearly the same time period between 2003 and 2012. In addition, Costa et al. (2017) reported that proxies of muscle power (i.e., standing long jump test) increased in 987 Portuguese adolescents aged 10 to 11 years between 1993 and 2013. In summary, these three selected original studies exemplify that national studies reported mixed results regarding secular trends for the same components of physical fitness over recent decades. This reemphasizes the need for a comprehensive study to update currently available systematic reviews and meta-analysis to provide an overview of the development of children’s and adolescents’ physical fitness status worldwide. For example, if children’s and adolescents’ physical fitness performance decreased over time, this alarming result can be used to design national or international initiatives in physical fitness promotion to ensure a holistic development and a lifelong healthy life.

2.2.2 Effects of age and sex

Physical fitness develops progressively during childhood and adolescence and is modulated by age and sex (De Miguel-Etayo et al., 2014; Golle et al., 2015; Malina et al., 2004; Niessner et al., 2020; Santos et al., 2014; Tambalis et al., 2016; Tomkinson et al., 2018). Accordingly, many studies worldwide (Castro-Piñero et al., 2010; Catley & Tomkinson, 2013; De Miguel-Etayo et al., 2014; Dobosz et al., 2015; Golle et al., 2015; Niessner et al., 2020; Ortega et al., 2011; Roriz de Oliveira et al., 2014; Santos et al., 2014; Tambalis et al., 2016; Tomkinson et al., 2018) revealed developmental gains in physical fitness from childhood to adolescence (De Miguel-Etayo et al., 2014; Dobosz et al., 2015; Golle et al., 2015; Niessner et al., 2020; Ortega et al., 2011; Santos et al., 2014; Tambalis et al., 2016). Boys show better performances compared to girls during childhood and adolescence in most components of physical fitness (Castro-Piñero et al., 2010; Catley & Tomkinson, 2013; De Miguel-Etayo et al., 2014; Dobosz et al., 2015;

Fraser et al., 2019; Lisowski et al., 2020; Ortega et al., 2011; Roriz de Oliveira et al., 2014; Santos et al., 2014; Tambalis et al., 2016; Tomkinson et al., 2018), except for balance (De Miguel-Etayo et al., 2014; Dobosz et al., 2015) and flexibility (Catley & Tomkinson, 2013; De Miguel-Etayo et al., 2014; Dobosz et al., 2015; Golle et al., 2015; Lisowski et al., 2020; Ortega et al., 2011; Roriz de Oliveira et al., 2014; Santos et al., 2014; Tambalis et al., 2016; Tomkinson et al., 2018).

However, when taking a closer look at the development of physical fitness components, sex- and age-differential developmental differences between the components occur. For instance, in a four-year longitudinal study including 240 children, Golle et al. (2015) highlighted sex- and age-differential developmental differences for different components of physical fitness in children aged 9 to 12 years. These components comprised of cardiorespiratory endurance (i.e., 9-min-run test), coordination (i.e., star-run test), speed (i.e., 50-m linear sprint test), proxies of lower (i.e., triple hop test) and upper (i.e., ball-push test) limbs muscle power. They reported that speed as well as proxies of lower and upper limbs muscle power were characterized by a mostly linear trajectory, whereas cardiorespiratory endurance and coordination followed a curvilinear development (Golle et al., 2015). These age-related differences are mostly due to growth (i.e., increasing body mass and body height) and maturation processes during childhood and adolescence (Malina et al., 2004). For instance, increasing body mass may have a positive effect on ball-push test performance but a negative effect on performance in tests such as the 9-min-run test which requires the constant acceleration of the body (Drenowatz et al., 2021). Furthermore, the fact that boys persistently showed better results than girls in all physical fitness tests (Golle et al., 2015) can be attributed to differences in muscle mass favoring boys, as there is evidence that prepuberal boys have on average 3.7% larger muscle mass than girls (Kanehisa et al., 1995; Malina et al., 2004). Consistent with this argument, Golle et al. (2015) reported larger sex differences in physical fitness tests requiring muscle mass (e.g., ball-push test) than tests of motor coordination (e.g., star-run test). This finding is supported by a study conducted by Overmann (2004). The author did not find any significant sex differences in cognitive tasks in children between three years of age and puberty.

The longitudinal study of Golle et al. (2015) and other above mentioned cross-sectional studies (Dobosz et al., 2015; Niessner et al., 2020; Ortega et al., 2011; Santos et al., 2014; Tambalis et al., 2016; Tomkinson et al., 2018) reported performance enhancements for all components of physical fitness during childhood and adolescence. All of these studies have in common that they examined a wide age range spanning childhood to adolescence (e.g., 9 to 17 years (Tomkinson et al., 2018)), that is a *long-term ontogenetic* perspective. However, a *short-term* *ontogenetic* perspective (i.e., focusing on a single year of life) might unveil different growth trajectories that might be overlooked within a *long-term ontogenetic* perspective. For example, the situation of boys' and girls' divergence during puberty might be different during prepuberty. Girls mature visibly about two years earlier than boys (Stratton & Oliver, 2020), but sex hormones rise already considerably starting at eight years of age (Marshall & Tanner, 1969, 1970; Rosenfield et al., 2009). If the early rise of sex hormones is related to body composition, the question arises whether there is evidence for faster development of girls than boys in this single year of life (i.e., ninth year), leading to a convergence of performance in physical fitness during the transition from pre-puberty to puberty. A *long-term ontogenetic* perspective on this question might be too 'long' to unveil any differences, therefore a *short-term ontogenetic* perspective might be the better approach.

A reanalysis of the longitudinal data of Golle et al. (2015) revealed that neither the *short-term ontogenetic* age differences nor the interactions with sex were significant within a single year of assessment (i.e., between 9.00 and 9.99 years old; see insets in Fig. 1 and Supplement B of Publication II). The only significant interaction between age and sex was found for proxies of upper limbs muscle power (see right panel in Fig. 1 of Publication II) suggesting divergence between boys and girls. However, the lack of evidence for significant *cross-sectional* age differences and their interactions with sex within a single year of life at a prepubertal age should not be taken as a proof for their lack (Altman & Bland, 1995), because statistical power may not have been strong enough. Yet, *cross-sectional* age differences and their interactions with sex within a single year of life at a prepubertal development stage have not been examined within a very large representative sample (> 100,000 children). With a large sample it is possible to zoom into a *cross-sectional short-term ontogenetic* window that might reveal new insights that can further the understanding of children's development of physical fitness. Furthermore, findings can be helpful for an individual evaluation of children's physical fitness status.

2.2.3 Effect of timing of school enrollment

The studies on the development of physical fitness during childhood and adolescence mentioned in chapter 2.2.2 were conducted in children and adolescents aged 5 to 18 years. All studies have in common that children and adolescents were divided into whole-year age groups. These whole-year age groups are also found in many areas of children's daily lives. For instance, children are divided into grades within schools or in 1-year age teams within sport clubs.

However, those whole-year age groups are not without limitations because the relative age varies depending on the specific cut-off date. Within schools, children are matched into grades based on their date of birth and a country-specific school enrollment date. For instance, in the Federal State of Brandenburg, Germany September 30th is the official date for school enrollment. Accordingly, children who are between 6.00 and 6.99 years old on September 30th are enrolled to school (i.e., keyage-children in first grade). This age grouping can result in an age-difference between children of nearly one year (Sandercock et al., 2013): children born on September 30th or slightly later are at the extreme end almost 1 year older than their classmates born later in August. These differences in birthdate can affect anthropometric characteristics (e.g., body height, body mass) as well as physical fitness (e.g., muscular strength, proxies of muscle power, cardiorespiratory endurance, or speed) (Cobley et al., 2009) because physical fitness performance develops with increasing age during childhood and adolescence (De Miguel-Etayo et al., 2014; Dobosz et al., 2015; Niessner et al., 2020; Ortega et al., 2011; Santos et al., 2014; Tambalis et al., 2016). Therefore, within a school grade, the relatively older ones (i.e., born near to the cut-off date) may show better performances than their relatively younger classmates (i.e., born later to the cut-off date) because of differences in relative age (Cobley et al., 2009; Musch & Grondin, 2001). In fact, the second publication of this doctoral thesis showed that physical fitness developed linearly with chronological age in a sample of 108,295 keyage third graders (i.e., children aged 8.00 to 8.99 years) (Fühner et al., 2021b). Moreover, the second publication indicated that even within the single ninth year of life, the relatively older classmates (i.e., aged 8.5 to 8.99 years) significantly showed better performances compared to their younger classmates (i.e., aged 8.0 to 8.49 years) in physical fitness (Fühner et al., 2021b).

However, within one school grade, there are also children who are younger (YTK) or older (OTK) than keyage-children because of early or late school enrollment, skipping or repetition of a school year. For example, due to individual developmental processes, parents have the option to enroll a gifted child earlier or a slow developing child later (Ministerium für Bildung Jugend und Sport, 2018). According to official statistics of the Federal State of Brandenburg (Germany), 1% of primary school children are enrolled earlier into school, i.e. they are younger than 6.00 years (Ministerium für Bildung Jugend und Sport, 2021a), 15% are enrolled later into school, i.e., they are older than 6.99 years (Ministerium für Bildung Jugend und Sport, 2021a), and 1% have to repeat a school year (Ministerium für Bildung Jugend und Sport, 2021b).

Given that there are already large differences in physical fitness within the group of keyage-children (Fühner et al., 2021b), the question arises how physical fitness of OTK and YTK children develops compared to keyage-children. However, the effect of timing of school enrollment on physical fitness development remains unclear and it has not yet been investigated how physical fitness of OTK and YTK children develops compared to keyage-children.

Studies from a related research topic analyzing academic performance of OTK and YTK children compared to keyage-children revealed large divergencies in academic performance (Jaekel et al., 2015; Martin, 2009; Urschitz et al., 2003). For instance, in a study including 1,144 German third graders, Urschitz et al. (2003) revealed that especially OTK children aged > 9.00 years showed worse academic performance (i.e., grades in mathematics, science, reading, spelling, and handwriting) than keyage-children. In a study including 3,684 Australian high school students aged 14 years, Martin (2009) revealed that YTK children showed significantly better academic performance (i.e., performance in literacy and numeracy) compared to keyage-children. However, this has not yet been analyzed for physical fitness. The findings can further the understanding of children´s physical fitness development and can be interesting for physical education teachers to individually evaluate student’s physical fitness considering their age, sex, and timing of school enrollment. According to these information physical education teachers can create a learning setting for each child based on his / her individual needs to guarantee a holistic development (Marta et al., 2012).

3. Research objectives and hypotheses

The relevance of physical fitness for children's and adolescents' health is indisputable and it is crucial to regularly assess and evaluate children's and adolescents' individual physical fitness status to detect potential negative health consequences in time. Physical fitness tests are easy-to-administer, reliable, and valid which is why they should be widely used to provide information on performance development and health status of children and adolescents (Pate, 1991).

Although the development of children's and adolescents' physical fitness has been extensively described and analyzed in the literature, still some questions remain unknown that will be addressed in the present doctoral thesis to complement the knowledge of children's and adolescents' physical fitness development. Those findings can be used to prevent potential deficits in physical fitness as early as possible and consequently ensure a holistic development and a lifelong healthy life.

Previous systematic reviews and meta-analyses have examined secular trends in children's and adolescents' physical fitness (Dooley et al., 2020; Tomkinson & Olds, 2007; Tomkinson, 2007; Tomkinson et al., 2019). However, considering that those analyses are by now 15 years old and that updates are available only to limited components of physical fitness, it is time to re-analyze the literature and examine secular trends for selected components of physical fitness (i.e., cardiorespiratory endurance, muscle strength, proxies of muscle power, and speed). Furthermore, the available studies on children's development of physical fitness as well as the effects of moderating variables such as age and sex (De Miguel-Etayo et al., 2014; Golle et al., 2015; Malina et al., 2004; Niessner et al., 2020; Santos et al., 2014; Tambalis et al., 2016; Tomkinson et al., 2018) have been investigated within a *long-term ontogenetic* perspective. However, the effects of age and sex in the transition from pre-puberty to puberty in the ninth year of life using a *short-term ontogenetic* perspective and the effect of timing of school enrollment on children's development of physical fitness have not been clearly identified. Therefore, the primary aims of this doctoral thesis were to update secular trend analysis in selected components of physical fitness in children and adolescents, to examine *short-term ontogenetic cross-sectional* developmental differences in children`s physical fitness, and to compare physical fitness of OTK and YTK children versus keyage-children to further understand children's development of physical fitness (Figure 1).

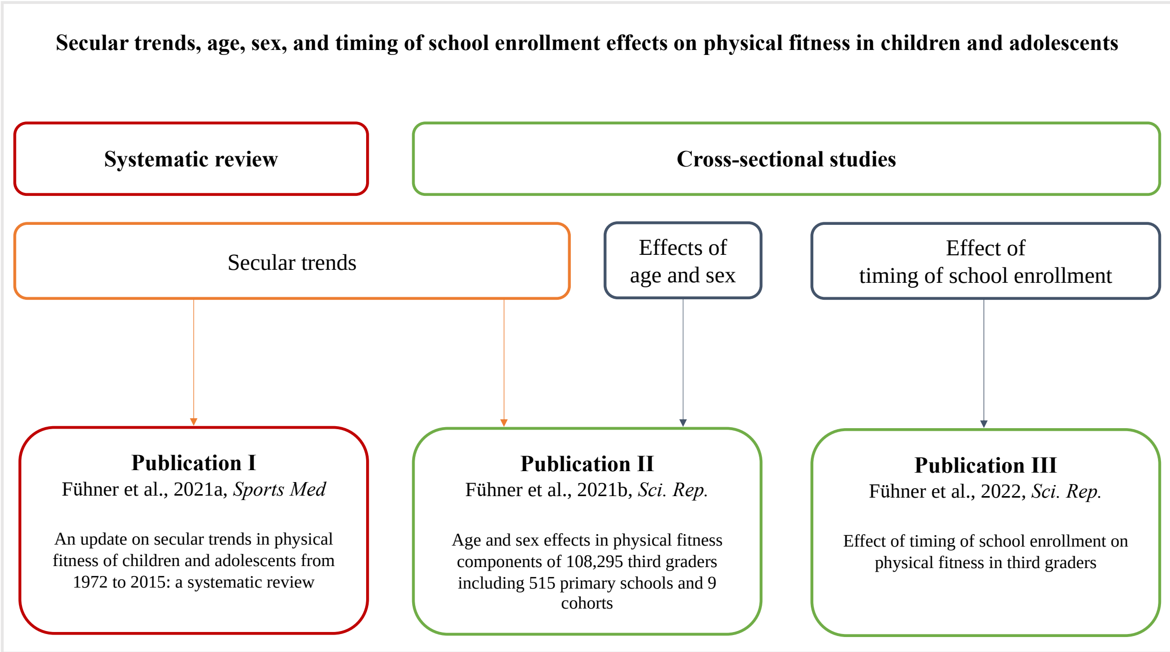

Figure 1 Schematic overview of the aims of the three publications (I, II, III) included in this doctoral thesis examining secular trends, age, sex, and timing of school enrollment effects on physical fitness in children and adolescents.

The research objectives and hypotheses of this doctoral thesis are based on the literature review in chapter 2 and can be summarized as follows.

Objective 1

The first objective was to provide an 'update' on secular trends in selected components of physical fitness (i.e., cardiorespiratory endurance, relative muscle strength, proxies of muscle power, speed) in children and adolescents aged 6 to 18 years.

Hypothesis 1

With reference to recently published studies (Dooley et al., 2020; Tomkinson et al., 2018) and due to the ongoing international promotion of physical activity and physical fitness (World Health Organization, 2018), a positive trend in physical fitness development of children and adolescents over time was hypothesized. In addition, it was expected that secular trends are heterogeneous in direction and magnitude and specific to the respective physical fitness component under consideration (see chapter 2.2.1).

Objective 2

The second objective was to quantify *short-term ontogenetic cross-sectional* developmental differences in physical fitness for five tests comprising health- and skill-related components of physical fitness in a large sample of 108,295 German eight-year old children.

Hypothesis 2

In this single year of life, it was expected to detect age differences that were not significant in the longitudinal study of Golle et al. (2015). Indeed, it was hypothesized that the component differences in developmental gains would be in general agreement with those in the longitudinal study of Golle et al. (2015) because gains should be larger (favoring older children) in tests that do not require continuous acceleration of the body as with the ball-push test vs the 9-min-run test (Drenowatz et al., 2021). Similarly, as in Golle et al. (2015) the test-related sex effects would be larger, the more muscle mass is required to perform a physical fitness test (Kanehisa et al., 1995; Malina et al., 2004) (favoring boys) and less for tests involving motor coordination (Overman, 2004) (see chapter 2.2.2).

Objective 3

Based on the cross-sectional study design (explained in objective two) with annual physical fitness assessments from 2011 to 2019, secular trends have to be taken into account as a source of variance in physical fitness data. Therefore, the third objective was to examine secular trends in five health- and skill-related components of physical fitness in a sample of 108,295 German eight-year old children.

Hypothesis 3

It was hypothesized that the results would be in line with recent systematic reviews and show a decline in cardiorespiratory endurance and an increase in speed (Fühner et al., 2021a; Tomkinson et al., 2019) (see chapter 2.2.1).

Objective 4

The fourth objective was to compare physical fitness of OTK and YTK children with keyage-children in a sample of German primary school children taking age, sex, school, and assessment year into account.

Hypothesis 4

With reference to the relevant school-based studies on differences in academic performance of OTK and YTK children versus keyage-children (Jaekel et al., 2015; Martin, 2009; Urschitz et al., 2003), it was hypothesized that OTK children would show poorer and YTK children better physical fitness compared with keyage-children (see chapter 2.2.3).

For this purpose, a systematic review and two cross-sectional studies were conducted and three manuscripts were prepared (Figure 1) and published in international peer-reviewed journals. The publications will be cited in this thesis with the following numbering:

Publication I Fühner, T., Kliegl, R., Arntz, F., Kriemler, S., & Granacher, U. (2021a). An update on secular trends in physical fitness of children and adolescents from 1972 to 2015: a systematic review. *Sports Medicine, 51*, 303-320. https://doi.org/10.1007/s40279-020-01373-x

The final publication is available at:

https://link.springer.com/article/10.1007/s40279-020-01373-x

Publication II Fühner, T., Granacher, U., Golle, K., & Kliegl, R. (2021b). Age and sex effects in physical fitness components of 108,295 third graders including 515 primary schools and 9 cohorts. *Scientific Reports, 11*, 17566. https://doi.org/10.1038/s41598-021-97000-4

The final publication is available at:

https://www.nature.com/articles/s41598-021-97000-4

Publication III Fühner, T., Granacher, U., Golle, K., & Kliegl, R. (2022). Effect of timing of school enrollment on physical fitness in third graders. *Scientific Reports, 12,* 7801. https://doi.org/10.1038/s41598-022-11710-x

The final publication is available at:

https://www.nature.com/articles/s41598-022-11710-x

4. Materials and methods

The following chapter briefly summarizes the participants' characteristics, the experimental procedures, the data collection and processing, and the statistical analyses. The applied methodological approach in this doctoral thesis was based on the deduced research hypotheses. Detailed information on materials and methods are provided in the respective sections of Publications I, II, and III.

4.1 Participants

The two cross-sectional studies (Publications II and III) were part of the ongoing EMOTIKON project (www.uni-potsdam.de/en/emotikon). Starting in 2009, physical fitness of all third graders living in the Federal State of Brandenburg, Germany is being tested annually (see Chapter 4.2).

The study sample in Publication II consisted of 108,295 healthy children from 515 different primary schools who had been enrolled within the legal key date of the Federal State of Brandenburg, that is in a given year of school enrollment they were at least 6.00 and at most 6.99 years old on September 30th and, therefore, varied between 8.00 and 8.99 years in the third grade (i.e., keyage-children). The study sample in Publication III focused on third graders who are younger (i.e., 7.00 to 7.99 years) or older (i.e., 9.00 to 9.99) than keyage-children. Younger children are enrolled in third grade for reasons such as early enrollment into school or if they skip a grade (i.e., younger-than-keyage [YTK] children). In contrast, older children are enrolled in third grade for reasons such as delayed enrollment into school or if they repeat a grade (i.e., older-than-keyage [OTK] children). Accordingly, 2,586 healthy YTK children aged 7.00 to 7.99 years and 26,540 healthy OTK children aged 9.00 to 9.99 years were included.

The two cross-sectional studies were mandated and approved by the Ministry of Education, Youth and Sport of the Federal State of Brandenburg, Germany. The Brandenburg School Law (Ministerium für Bildung Jugend und Sport, 2018) requires that parents are comprehensively informed prior to the start of the study. Consent was however not needed as the tests are obligatory for both, children and schools. Information in detail regarding the participants and participants' exclusion criteria can be found in the "Methods" section of the respective published manuscripts (see Publications II and III).

4.2 Experimental procedures

In Publication I, a systematic review on secular trends in selected components of physical fitness in children and adolescents was conducted. To ensure a standardized process, the systematic review was conducted in accordance with the Preferred Reporting Items for Systematic Reviews and Meta-Analysis (PRISMA) statement guidelines (Liberati et al., 2009; Moher et al., 2009). Studies identified through a systematic computerized literature search in the electronic databases *PubMed* and *Web of Science* were included in the systematic review if they provided relevant information with regards to a modified PICOS approach (Liberati et al., 2009). The considered parameters were population, outcomes, and study design.

Physical fitness tests for the two cross-sectional studies were administered to all third graders in the Federal State of Brandenburg, Germany annually between September and November from 2011 to 2019. Physical fitness was assessed with the EMOTIKON test battery (see chapter 4.3.1).

Physical fitness tests were administered by qualified physical education teachers of each school according to standardized test protocols during the regular physical education classes in the participating primary schools. Prior to testing, all third graders completed a standardized warm-up program consisting of various running exercises (e.g., side-steps) and small games (e.g., playing tag). Detailed information on the experimental procedures is presented in the "Methods" section of the respective Publications I, II, and III.

4.3 Data collection and processing

Both cross-sectional studies (Publications II and III) were conducted in the respective primary schools of the Federal State of Brandenburg, Germany. Measurements comprised the assessment of physical fitness.

4.3.1 Systematic literature search

A systematic computerized literature search was conducted in the electronic databases *PubMed* and *Web of Science* to identify studies that explicitly examined secular trends in physical fitness in healthy children and adolescents. To locate studies about secular trends in children's and adolescents' physical fitness, a Boolean search strategy was used. The detailed search strategy can be found in the "Methods" section of the Publications I. In addition, reference lists of each article as well as relevant review articles and meta-analyses were cross-referenced / screened to locate further suitable references to be considered in the systematic review. Studies were integrated in the systematic review if they gave relevant information with regards to a modified PICOS approach (Liberati et al., 2009). The considered parameters were population, outcomes, and study design. A detailed list of the predefined inclusion and exclusion criteria can be found in the "Methods" section of the Publication I. All identified studies were coded for the variables: nationality, physical fitness components, chronological age, sex, and trends. The relevant data were extracted in an Excel spreadsheet from all included studies. If studies included several tests of the same physical fitness component, only one test was used. Detailed information on the coding of studies is presented in Publication I.

4.3.2 Assessment of physical fitness

Physical fitness was assessed through the EMOTIKON test battery. The five tests measured cardiorespiratory endurance (i.e., 6-min-run test), coordination (i.e., star-run test), speed (i.e., 20-m linear sprint test), and proxies of lower (powerLOW [i.e., standing long jump test]) and upper limbs (powerUP [i.e., ball-push test]) muscle power. The EMOTIKON test battery officially comprises of six tests. In 2016, the assessment of flexibility (i.e., stand-and-reach test) was replaced by the assessment of balance (i.e., single-leg balance test with eyes closed). However, due to the much smaller number of scores and their confound with assessment year (Augste & Künzell, 2014; Hjorth et al., 2013) these two tests were not included in both cross-sectional studies. Detailed information on the assessment of each physical fitness component is presented in the respective "Methods" sections of the Publications II and III.

Cardiorespiratory endurance

Cardiorespiratory endurance was evaluated using the 6-min-run test. Participating children had to run the furthest possible distance during the six minutes test time around a volleyball field (i.e., 54 m) at a self-paced velocity. After the six minutes, maximal distance covered in meters was recorded. The 6-min-run test was reliable (test-retest) for children aged 7 to 11 years with an intraclass correlation coefficient (ICC) of 0.92 (Bös, 2009).

Coordination

Coordination under time pressure was assessed using the star-run test. During the star-run test, the participating children had to complete a parkour using different running techniques (i.e., running forwards, running backwards, side-steps) as quickly as possible. A detailed description of the star-run test can be found in the respective "Methods" sections of the Publications II and III. Time for test completion in seconds was taken using a stopwatch. The star-run test was reliable (test-retest) for children aged 8 to 10 years with an ICC of 0.68 (Schulz, 2013).

Speed

Speed was evaluated using the 20-m linear sprint test. After an acoustic signal, the participating children had to sprint as fast as possible over a distance of 20 m. Time for test completion was taken using a stopwatch. High test-retest reliability was reported for the 20-m linear sprint test with an ICC of 0.90 in children aged 7 to 11 years (Bös, 2009).

Proxies of lower limbs muscle power (PowerLOW)

PowerLOW was tested through the standing long jump test. The participating children had to jump as far as possible from a frontal position. Jump distance was recorded using a measuring tape. Test-retest reliability has been reported to be high for children aged 6 to 12 years with an ICC of 0.94 (Fernandez-Santos et al., 2015).

Proxies of upper limbs muscle power (PowerUP)

Power up was assessed with the ball-push test. From a standing position, the participating children had to push a 1 kg medicine ball that was held tight right in front of the chest. The pushing distance was recorded using a measuring tape. The ball-push test showed high test-retest reliability for children aged 8 to 10 years with an ICC of 0.81 (Schulz, 2013).

4.4 Statistical analyses

In all three publications statistical inference was based on Linear Mixed Models (LMM). LMM as well as pre- and post-processing of data were carried out in the *R* environment of statistical computing (R Core Team, 2020) and in the *Julia* programming language (Bezanson et al., 2017). Detailed information on pre- and post-processing of data, the *R* and *Julia* packages, and the model selection processes can be found in the respective "Methods" sections as well as in the supplementary materials of Publications I, II, and III. The following sections briefly summarize the final LMMs of each publication.

Publication I

To provide an 'update' on secular trends in selected components of physical fitness, a LMM estimated the effects of sex and linear, quadratic, and cubic secular trends nested under the four components of physical fitness (i.e., $8 \times 4 = 32$ fixed effects). For each observation, year of assessment was centered at 1990 (i.e., at the midpoint of first-year assessments between 1974 and 2006 across studies). The nested model specification denoted that for each of the four physical fitness components, the main effect of sex, the secular trend (i.e., three parameters for linear, quadratic, and cubic component), and the interactions between sex and secular trend were estimated. The LMM included the 22 studies of the systematic review as levels of a random factor. The final model considered study-related variance components (VCs) for cardiorespiratory endurance, relative muscle strength, sex, and linear yearly trends for cardiorespiratory endurance, relative muscle strength, and proxies of muscle power. With this random-effect structure, none of the high-order interactions between sex and quadratic or cubic secular trends were significant. Therefore, these eight fixed effects were removed from the LMM. The final LMM estimated 24 fixed effects, six VCs, and the residual variance. A |z|-value > 2.0 was adopted as a criterion for rejection of the null hypothesis.

Publication II

To examine *short-term ontogenetic cross-sectional* developmental differences in physical fitness, a LMM was used for statistical inference. As fixed effects, four sequential-difference contrasts for the physical fitness components were specified: (H1) coordination vs. cardiorespiratory endurance, (H2) speed vs. coordination, (H3) powerLOW vs. speed, and (H4) powerUP vs. powerLOW. In addition, the effect of age (centered at 8.50 years) as a second-order polynomial trend, the effect of sex, and all interactions between contrasts, age, and sex were considered as fixed effects. Considering the large number of observations, children, and schools, a two-sided |z|-value > 3.0 was used as significance criterion for interpretating the fixed effects.

The final LMM considered the three random factors child (N = 108,295), school (N = 515), and assessment year (N = 9). The random effect structure included: Grand Mean (varying intercepts) for all three random factors; component-related VCs and correlation parameters (CPs) for child; component-, age, and sex-related VCs and CPs for school; and component-related VCs for assessment year. The final LMM was also estimated with two alternative parameterizations that did not change the goodness of fit but provided information about CPs between test scores instead of test effects (i.e., contrasts).

Publication III

To compare physical fitness development of YTK and OTK children with that of keyage-children, the physical fitness performance for ages 7.00 to 7.99 and 9.00 to 9.99 years was predicted using parameter estimates of the final LMM used in Publication II. Through this predication analyses information on physical fitness performance of keyage-children at the ages 7.00 to 7.99 years and 9.00 to 9.99 years was obtained.

Through physical fitness testing in EMOTIKON, the actual physical fitness status of YTK children and OTK children was obtained, that is the children's observed physical fitness z-scores. The difference between observed and predicted z-scores provided a *delta z-score* that was used as dependent variable in the following LMMs to compare physical fitness development of YTK and OTK children (i.e., obtained scores through physical fitness testing) with that of keyage-children (i.e., predicted data).

The data were analyzed with separate LMMs for OTK and YTK children. For both LMMs, the four sequential-difference contrasts for the physical fitness components used in Publication II were included. The LMM for OTK children also considered the effect of age (centered at 8.50 years) as a second-order polynomial trend, the effect of sex, and all interactions between contrasts, age, and sex as fixed effects. A two-sided $|z|$-value > 2.0 was adopted as significance criterion for the interpretation of fixed effects.

The final LMM for OTK children considered the three random factors child (N = 26,540), school (N = 513), and assessment year (N = 9). The random effect structure included VCs and CPs between *delta z-scores* of the five components for child, VCs between *delta z-scores* of the five components for school, age-related VCs for school, and a Grand Mean (varying intercept) for assessment year. The random-effect structure of the parsimonious LMM of delta z-scores was expected to be simpler than the one for the LMM of the reference population in Publication II because of the much smaller number of children and because many of the effects, VCs and CPs were already part of the predicted z-scores.

Because of the relatively small sample size of YTK children (N = 2,586) compared to OTK children (N = 26,540), the complexity of the final LMM for YTK children was considerably reduced. The final LMM for YTK children included child (N = 2,586), school (N = 437), and assessment year (N = 9) as three random factors. The random effect structure included VCs and CPs between *delta z-scores* of the five components for child, and VCs between *delta z-scores* of the five components for school.

5. Results

The following chapter provides a brief summary of the main results of Publications I, II, and III included in the present doctoral thesis. The results in detail are provided in the respective Publications I, II, and III.

5.1 Secular trends in physical fitness

After the search and selection process, 22 studies were finally identified as eligible and included in the systematic review. Secular trends in selected components of physical fitness (i.e., cardiorespiratory endurance, relative muscle strength, proxies of muscle power, speed) were examined for N = 96,522 children and adolescents aged 6 to 18 years from 17 mainly high-income countries in the United States, Australia, and Europe. The observation period was between 1972 and 2015.

For cardiorespiratory endurance, an initial increase with a local maximum in 1986 was followed by a decline to a local minimum in 2010 with a negative inflection point in the turn of the century. Since 2010, the decline in cardiorespiratory endurance has diminished and performances tended to increase slightly again (i.e., positive cubic secular trend; $b = 0.00012$, $z = 3.94$). There was a steeper linear decrease in cardiorespiratory endurance for boys compared to girls ($b = - 0.00796$, $z = - 3.28$). The magnitude of secular trend in cardiorespiratory endurance was large, approximately 1.5 standard deviations in both sexes.

Relative muscle strength declined until 1982 and then increased until 2006 with a positive inflection point in 1990 (i.e., cubic secular trend; $b = - 0.00004$, $z = - 2.14$). Although there were no significant effects of sex, the trend was more pronounced for boys. Regarding the effect size, relative muscle strength showed small secular trends with a narrow range of 0.50 standard deviations in both sexes.

For proxies of muscle power, an initial linear increase with a local maximum in 1982 for girls and 1994 for boys ($b = 0.00474$, $z = 2.80$) was followed by a negative quadratic decline ($b = - 0.00060$, $z = - 2.04$). The magnitude of secular trend in proxies of muscle power was small-to-medium, approximately 0.5 standard deviations in both sexes.

For speed, boys and girls showed similar secular trends. The shape of the function was characterized by a significant negative linear ($b = - 0.01822$, $z = - 4.90$) and a significant positive cubic ($b = 0.00006$, $z = 4.60$) secular trend. An initial increase with a local maximum in 1980 was followed by a decline to a local minimum in 2002 with a negative inflection point in 1990.

Since then, speed has been increasing again. With 0.5, the effect sizes for boys and girls are small.

Results on characteristics of the included studies, conditional means, and observation-level model residuals are described in detail in the results section of Publication I.

5.2 Effects of age and sex on physical fitness

Detailed information on the effects of age and sex on physical fitness development and goodness-of-fit statistics as well as model residuals are provided in the 'Results' sections of Publication II.

Main effects of age

Findings revealed that despite the small age range, the age-related differences were large and evident for all physical fitness components. The overall linear trend for age was significant (b = 0.271, z = 31.60). The age effect was larger for coordination than cardiorespiratory endurance (H1; b = 0.216, z = 18.44), larger for coordination than speed (H2; b = - 0.069, z = - 6.22), and larger for powerUP than powerLOW (H4; b = 0.307, z = 25.87). The difference between age slopes for powerLOW and speed was not significant (H3; b = - 0.007, z = - 0.69).

Main effects of sex

Findings revealed that the sex-related differences were large and evident for all physical fitness components, with boys outperforming girls. The overall sex effect (averaged over age) was significant (b = 0.413, z = 86.52). The sex effect was larger for cardiorespiratory endurance than coordination (H1; b = - 0.260, z = - 39.03), larger for speed than coordination (H2; b = 0.078, z = 12.35), larger for powerLOW than speed (H3; b = 0.066, z = 11.24), and larger for powerUP than powerLOW (H4; b = 0.296, z = 43.84).

Interaction effects of age and sex

There was no evidence for a significant age x sex interaction or significant age x sex x component interactions (all $b \leq 0.039$, $|z| \leq 1.74$).

Variance components (VCs)

The component-related VCs were large for children (0.72 to 0.78), medium for schools (0.24 to 0.37), and small for assessment year (0.03 to 0.06). VCs for the age-related gains (0.10) and the sex effect (0.05) were also small for schools.

VCs for component contrasts were large (0.47 to 0.72) for children and somewhat smaller, but still highly reliable (0.30 to 0.40) for schools, especially in contrast to VCs estimated for school-related age (0.10) and sex (0.05) effects, and especially in contrast to assessment-year-related effects (0.04 to 0.09).

Assessment-year-related variance components (VCs)

The small, but significant assessment-year-related VCs showed that there were reliable component x assessment year interactions. Findings revealed that performance in cardiorespiratory endurance declined and performance in speed increased between 2011 and 2019. In contrast, the initial increase in coordination, powerLOW, and powerUP was followed by a decrease in recent years.

5.3 Effect of timing of school enrollment on physical fitness

Detailed information on the effect of timing of school enrollment on physical fitness development are provided in the 'Results' sections of Publication III.

Older-than-keyage (OTK) children

Findings revealed a significant overall negative linear trend for age (b = -1.014, z = - 6.68) and a significant positive quadratic trend of age (b = 0.357, z = 4.50). The main effect of H1 was significantly larger for coordination than for cardiorespiratory endurance (b = 0.230, z = 2.51). The linear age developmental rate (averaged across sex) was larger for cardiorespiratory endurance compared to coordination (H1; b = - 0.693, z = - 3.38) and the quadratic age developmental rate was larger for coordination compared to cardiorespiratory endurance (H1; b = 0.294, z = 2.73).

Three of the four component contrasts interacted with sex: the *delta z-score* was more negative for boys than girls with regards to cardiorespiratory endurance and more negative for girls than boys regarding coordination (H1; b = 0.052, z = 3.41), the negative difference between boys and girls in the delta z-score was larger for powerLOW than speed (H3; b = 0.042, z = 3.38), the same powerLOW sex difference was the source of the significant interaction for the fourth contrast (H4; b = - 0.044, z = - 2.80).

The VCs between *delta z-scores* were large for children (0.88 to 0.98) and small for schools (0.06 to 0.09). VCs for the age-related gains (0.09) was also small for schools.

Younger-than-keyage (YTK) children

The Grand Mean was significant (b = 0.082, z = 5.09). The main effect was larger for coordination than for cardiorespiratory endurance (H1; b = 0.074, z = 3.05), larger for coordination than for speed (H2; b = - 0.055, z = - 2.49) and larger for powerUP than for powerLOW (H4; b = 0.060, z = 2.47).

The VCs between *delta z-scores* were large for children (0.79 to 0.89) and small for schools (0.09 to 0.12).

6. General discussion

In the present cumulative doctoral thesis, findings obtained from a systematic review (Publication I) and two cross-sectional studies (Publication II and III) were incorporated to complement current knowledge on children's development of physical fitness. More precisely, the aims of the thesis were to update secular trends in selected components of physical fitness in children and adolescents aged 6 to 18 years, to address *short-term ontogenetic cross-sectional* developmental differences in physical fitness in a large sample of eight-year old children, and to compare physical fitness of OTK and YTK children versus keyage-children in a sample of German primary school children.

Overall, results of the systematic review revealed a large initial improvement and an equally large subsequent decline between 1986 and 2010 as well as a stabilization between 2010 and 2015 in cardiorespiratory endurance, a general trend towards a small improvement in relative muscle strength, an overall small negative quadratic trend for proxies of muscle power, and a small-to-medium improvement in speed since 2002. Findings from the cross-sectional studies showed that even in a single prepubertal year of life physical fitness performance developed linearly with increasing chronological age, boys showed better performances than girls in all physical fitness components, the components varied in the size of sex and age effects, and secular trends varied between physical fitness components. Furthermore, findings from the cross-sectional studies revealed that OTK children showed poorer performance compared to keyage-children, especially in coordination, OTK girls showed better performances than OTK boys, and YTK children outperformed keyage-children, especially in coordination. The following sections provide a brief interpretation and discussion of the main findings based on the available body of literature.

6.1 Secular trends in physical fitness

Contrary to hypothesis 1, the systematic review revealed a decline in cardiorespiratory endurance, particularly between 1986 and 2010. However, it has to be noted that performance tended to stabilize between 2010 and 2015. These findings complement a recently published analysis about secular trends in cardiorespiratory endurance by Tomkinson et al. (2019). They reported that the international rate of decline in cardiorespiratory endurance has diminished and stabilized since 2000 (Tomkinson et al., 2019). Moreover, Tomkinson et al. (2019) showed that decreases in cardiorespiratory endurance were smaller for girls compared to boys which is also confirmed by the results of the systematic review. These consistent results between the systematic review and the analysis of Tomkinson et al. (2019) underline that the two analyses confirm each other despite their different methods and statistical approaches. While Tomkinson et al. (2019) included data for the 20-m shuttle run, the systematic review additionally considered data for time- and distance-related cardiorespiratory endurance tests. Tomkinson et al. (2019) estimated the $\dot{V}O_{2peak}$ from 20-m shuttle run performance for children and adolescents aged 8 to 19 years, irrespective of sex by using the equation generated by Léger et al. (1988).

Tomkinson and Olds (2007) stated that secular trends in physical fitness were modulated by several factors such as altered social, behavioral, physical, psychosocial, and physiological aspects. Specifically, psychosocial aspects such as motivation, willingness to exert maximal effort, and pacing strategies may have affected the performance, but unfortunately these aspects cannot be controlled. According to Olds and Dollman (2004) and Albon et al. (2010), physiological aspects such as BMI and percentage of body fat mass have the greatest impact on secular trends in cardiorespiratory endurance. They stated that changes in BMI affected changes in cardiorespiratory endurance between 40% (Olds & Dollman, 2004) to 70% (Albon et al., 2010) which seems logical considering that during running the increased body and / or fat mass must be permanently accelerated (Tomkinson & Olds, 2007). In this context, Cureton et al. (1978) reported that an increase in body and / or fat mass increased the energy expenditure during running resulting in a comparatively low relative $\dot{V}O_{2max}$. Interestingly, the NCD Risk Factor Collaboration reported an increased prevalence of overweight and obesity in children and adolescents worldwide between 1975 and 2016 but noted that the trend has slightly declined recently. However, the prevalence is still at a high level, especially in high-income countries (Bentham et al., 2017).

Another aspect that might have an impact on secular trends in cardiorespiratory endurance are the low levels of physical activity. A recently published analysis by Guthold et al. (2020) revealed that in 2016 only 20% of children and adolescents (girls: 15.3% and boys: 22.4%) worldwide met the WHO recommendations of 60 min of moderate-to-vigorous daily physical activity (World Health Organization, 2010). Compared to data from 2001, the low levels of physical activity among girls have remained relatively stable over the past 15 years (14.9%), while they have declined slightly among boys (19.9%). Given that physical activity and physical fitness are interdependent (Larsen et al., 2015), the low(er) levels of physical activity might lead to a reduced training stimulus and consequently to diminished cardiorespiratory endurance (Olds & Dollman, 2004).

Contrary to hypothesis 1, findings from the systematic review revealed for relative muscle strength, an initial decline with a local minimum in 1982 that was followed by an increase to a local maximum in 2006 before declining again. These findings complement recently published analyses about secular trends in relative (Kaster et al., 2020) and absolute (Dooley et al., 2020) muscle strength. Kaster et al. (2020) used solely data on the sit-up test of 9,939,289 children and adolescents aged 9 to 17 years living in 31 countries. They showed that the international rate of muscle strength improvement has slowed between 1964 to 2000, then stabilized until 2010, before declining again (Kaster et al., 2020). Dooley et al. (2020) showed that the international rate of increase in handgrip strength progressively improved between 1967 and 2017 in 2,216,320 children and adolescents aged 9 to 17 years from 19 countries.

Contrary to hypothesis 1, findings from the systematic review revealed a small negative quadratic trend with different peaks in 1982 for girls and 1994 for boys in proxies of muscle power. These findings complement a recently published analysis about secular trends in proxies of muscle power using solely data on the standing long jump test for 10,940,801 children and adolescents aged 9 to 17 years living in 29 countries (Tomkinson et al., 2021). They showed that the international rate of proxies of muscle power improvement between 1960s and 1980s has slowed in the 1990s, before declining (Tomkinson et al., 2021). One factor that coincides with the recent declines in relative muscle strength and proxies of muscle power is the small percentage of European adolescents aged 15 to 17 years (♂ = 24.7%; ♀ = 13.8%) who meet the WHO recommended level of muscle-strengthening activity at three times a week (Bennie et al., 2021). Thus, the low(er) levels of muscle-strengthening activity might lead to a diminished exercise stimulus and consequently to reduced relative muscle strength and proxies of muscle power. Faigenbaum and MacDonald (2017) used the term *paediatric dynapenia* to characterize children and adolescents who have low levels of muscle strength and proxies of muscle power. However, they emphasized that the low levels were a changeable risk factor for preventing lifelong health-related diseases through appropriate interventions (Faigenbaum & MacDonald, 2017). The design of appropriate interventions will be further discussed in chapter 8.

In line with hypothesis 1, findings from the systematic review indicated small-to-medium improvements in recent years with regards to speed. Because this systematic review was the only work to date that updated secular trends for speed, results are compared with an earlier comprehensive analysis conducted by Tomkinson (2007). Tomkinson (2007) summarized the existing literature on secular trends regarding speed for the timespan between 1958 and 2003. The analysis considered physical fitness data of 28,320,308 children and adolescents aged 6 to 19 years from around the world. The findings revealed that speed increased by + 0.04% p.a.

between 1958 and 2003. However, a detailed analysis showed that between 1958 and 1985 speed initially increased by + 0.27% p.a. before declining by - 0.08% p.a. (Tomkinson, 2007). The systematic review showed the same results for the mentioned time spans: an initial increase with a local maximum in around 1980 was followed by a decline to a local minimum in 2002. Consequently, the two studies corroborate each other, despite different methodological approaches, such as different data sets and number of studies included. The systematic review now adds to the overall research that speed has been increasing again since 2002.

Compared to secular trends in relative muscle strength, proxies of muscle power, and speed (0.5 to 1.0 SD), larger trends were examined for cardiorespiratory endurance (1.5 SD for boys and for girls). One explanation for the varying effects between components of physical fitness might be the different influence of BMI / percentage of fat mass. Tomkinson et al. (2007) analyzed to what extent BMI / percentage of fat mass affects proxies of muscle power, speed, and cardiorespiratory endurance. The correlations between BMI / percentage of fat mass and proxies of muscle power (i.e., standing long jump, $r = -0.20$) or speed (i.e., 50-m sprint, $r = -0.24$) were distinctively smaller than the correlations between BMI / percentage of fat mass and cardiorespiratory endurance (i.e., 1600 m run, $r = -0.52$) (Tomkinson & Olds, 2007). In addition, Tomkinson et al. (2007) revealed that for fat-free mass the effect was reversed. The correlations between fat-free mass and cardiorespiratory endurance ($r = 0.21$) were smaller than for proxies of muscle power ($r = 0.59$) and speed ($r = 0.57$). Based on the results of the systematic review and the evidence available within the literature, it remains unknown whether this is an issue of fatness or fitness. Considering the analyses by Tomkinson et al. (2007), it appears that fatness affects cardiorespiratory endurance in particular.

In summary, the systematic review revealed that there was no generic secular trend across the different physical fitness components and years, confirming hypothesis 1. This finding was also supported by the results of the cross-sectional study (Publication II). In line with hypothesis 3, all considered physical fitness components showed different patterns of secular trends across the nine years from 2011 to 2019. Performance in cardiorespiratory endurance declined and performance in speed increased between 2011 and 2019. In contrast, the initial increase in coordination, powerLOW, and powerUP was followed by a decrease in recent years. The reported secular trends for cardiorespiratory endurance, speed, and proxies of muscle power confirm the results of recently published systematic reviews (Fühner et al., 2021a; Tomkinson et al., 2019). However, there is only limited evidence for secular trends in coordination. In a recently published review, Eberhardt et al. (2020) could only identify three studies that examined secular trends in coordination with each study reporting conflicting results regarding increase, decline, and stagnation of coordination over time. They stated that the inconsistency between the studies was the result of the complex and multidimensional construct of coordination with different coordination tests measuring different aspects of coordination (Eberhardt et al., 2020). Therefore, it is difficult to compare the results of the cross-sectional study with other studies and more research is needed to further understand secular trends in coordination.

Overall, the systematic review (Publication I) and the cross-sectional study (Publication II) revealed that there was no generic secular trend across the different physical fitness components and years. Therefore, physical fitness components should be analyzed separately because each component exhibited a different pattern of secular trend.

6.2 Effects of age and sex on physical fitness

Well in line with hypothesis 2, boys showed better performance than girls in all four physical fitness components (i.e., cardiorespiratory endurance, coordination, speed, power [LOW / UP]). This finding corresponds with several studies providing physical fitness normative values (Catley & Tomkinson, 2013; De Miguel-Etayo et al., 2014; Ortega et al., 2011; Ramos-Sepúlveda et al., 2016; Roriz de Oliveira et al., 2014; Santos et al., 2014; Tambalis et al., 2016; Tomkinson et al., 2018; Woll et al., 2011). For instance, Tambalis et al. (2016) stated that boys aged 6 to 18 years significantly outperformed girls in cardiorespiratory endurance (i.e., 20-m shuttle run test), powerLOW (i.e., standing long jump test), and agility (i.e., 10 × 5 m agility shuttle run test). In addition, De Miguel-Etayo et al. (2014) reported that boys aged 6 to 10 years showed significantly better performances for cardiorespiratory endurance (i.e., 20-m shuttle run test), speed (i.e., 40-m linear sprint test), and powerLOW (i.e., standing long jump test) compared to girls. Well in line with hypothesis 2, the components varied in the size of sex effects. These differences in sex-related effects can mostly be explained by differences in body composition between boys and girls. For example, the smaller sex effect in powerLOW compared to powerUP can be attributed to a poorer proportion of strength relative to body mass especially in the upper limbs in girls compared to boys (Beunen & Thomis, 2000; Round et al., 1999). The large difference in cardiorespiratory endurance can be attributed to physiological factors such as girls' poorer mechanical efficiency and fractional utilization of oxygen compared to boys (Armstrong & Welsman, 2007; Tomkinson et al., 2018). In addition, the smaller muscle mass (Malina et al., 2004) and smaller muscle cross-sectional area (Kanehisa et al., 1995) of girls compared to boys limit girls' performances in physical fitness components that require muscle mass. The sex effect was also significantly smaller for speed than for powerLOW. Performances in PowerLOW require a high proportion of muscle mass where boys usually show better performances than girls (De Miguel-Etayo et al., 2014; Tambalis et al., 2016; Tomkinson et al., 2018). In contrast, speed is less affected by muscle mass and more by motor coordination where sex differences were comparatively small (Ardila et al., 2011) or were not found at all (Overman, 2004). Therefore, the sex effect in speed might be smaller than in powerLOW. Obviously, the demand of motor coordination relative to proxies of muscle power and cardiorespiratory endurance is larger in coordination than in speed (Koutsandréou et al., 2016; Ludyga et al., 2020; Schmidt et al., 2017) and this could be one explanation why the sex effect was larger for speed than for coordination. The observed decrease of sex differences in performances with increased demands for motor coordination appears to be plausible. Coordination is part of many different central (e.g., brain) and peripheral (e.g., motor units) regions within the nervous system (Viru et al., 1999). The less a test engages the brain, the more sex is a relevant performance limiting factor. In summary, well in line with hypothesis 2, the increase of size of sex effects across tests can most likely be attributed to an increase in the relevance of muscle mass (favoring boys over girls) (Malina et al., 2004) and a decrease in the relevance in motor coordination (associated with small or no sex effects) (Ardila et al., 2011; Overman, 2004).

In line with hypothesis 2, older children showed significantly better performances for all physical fitness components compared to younger children within the ninth year of life. However, the components varied in the size of age effects. The different age-related gains for tests correspond to the current knowledge about the development of anthropometric parameters: increasing body mass and body height positively impact the performance in skill-related physical fitness components of proxies of muscle power, coordination, and speed. In contrast, the growth in body mass has a negative influence on weight-bearing performance in tests such as cardiorespiratory endurance (e.g., 6-min-run test) where the own body mass must be accelerated permanently. The larger age effect for skill-related components (i.e., speed, coordination, proxies of muscle power) compared to cardiorespiratory endurance corresponds to a study conducted by de Miguel-Etayo et al. (2014). They found significant age differences in powerLOW (i.e., standing long jump test) and strength (i.e., absolute handgrip strength test) but could not detect age differences in cardiorespiratory endurance (i.e., 20-m shuttle run test) in a sample of children aged 6 to 10 years. In line with this, Viru et al. (1998) reported that an accelerated increase in cardiorespiratory endurance appeared at later ages (i.e., 11 to 15 years in boys and 11 to 13 years in girls). The three components coordination, speed, and powerLOW have the relevance of proxies of muscle power in common, but they vary in the relevance of coordination. As discussed above, the sex effect within these three tests (i.e., star-run test < 20-m linear sprint test < standing long jump test) corresponds to their ranking on motor coordination. The age effect within these three tests was star-run test > 20-m linear sprint test and standing long jump test, which is in line (roughly) with their ranking on motor coordination. Therefore, well in line with hypothesis 2, the more a test engages the brain, the larger is the effect of age. The only test that did not fit into this scheme was the ball-push test assessing powerUP. The ball-push tests showed the largest age effects. The special status of powerUP will be discussed in the next section.

Apparently, performance in powerUP was affected by aspects other than physical fitness. One explanation could be that body mass has a strong influence on the performance in tests that measure powerUP. Heavier children generally have a higher percentage of muscle mass than normal weighted children (Ducher et al., 2009). While a higher percentage of muscle mass positively affects performances in non-weight-bearing tests (e.g., ball-push test), it negatively impacts performances in weight-bearing tests (e.g., standing long jump test) because the body mass has to be permanently accelerated compared to non-weight-bearing tests (Drenowatz et al., 2021). Thus, tests assessing powerUP favor heavier children for whom strength of upper limbs and upper body may lead to a limited performance in weight-bearing tests assessing cardiorespiratory endurance or coordination. Therefore, a good performance in powerUP may be more suggestive of a lack of overall physical fitness.

To sum up, the *short-term ontogenetic* results of the cross-sectional study (Publication II) revealed component-specific age and sex effects. The findings imply that physical education teachers, coaches, or researchers can utilize a proportional adjustment to individually interpret physical fitness of prepubertal school-aged children. In addition, physical education teachers, coaches, or researchers should be cautious in the evaluation of the ball-push test and should take into account that it might not assess physical fitness to the same degree as the other four tests.

6.3 Effect of timing of school enrollment on physical fitness

Well in line with hypothesis 4, the cross-sectional study revealed that OTK children exhibited poorer performance compared to keyage-children and YTK children outperformed keyage-children. Many studies reported a linear increase in physical fitness performance with increasing age (De Miguel-Etayo et al., 2014; Santos et al., 2014; Tambalis et al., 2016;

Tomkinson et al., 2018). For instance, in a study with 424,328 Greek children and adolescents aged 6 to 18 years, Tambalis et al. (2016) observed a linear devolvement in physical fitness performance with increasing age for cardiorespiratory endurance (i.e., 20-m shuttle run test), muscular strength (i.e., sit-up test), agility (i.e., 10 × 5 m shuttle run test), proxies of lower limbs muscle power (i.e., standing long jump test), and flexibility (i.e., sit-and-reach test). The increase of physical fitness in keyage-children corresponds to the findings of former studies. For keyage-children, physical fitness performance developed linearly with increasing age. However, the development of physical fitness for OTK children showed a different pattern. Poorer performance was assessed in OTK children aged 9.00 to 9.99 years in comparison to age-matched keyage-children for cardiorespiratory endurance, speed, proxies of muscle power, and especially for coordination. The reason for this deviation could be that third graders aged 9.00 to 9.99 years (i.e., OTK children) are not representative for the "average" age-matched keyage-child which is why a deviation from the typically reported physical fitness development with increasing age in this cohort was observed (De Miguel-Etayo et al., 2014; Santos et al., 2014; Tambalis et al., 2016; Tomkinson et al., 2018). The exact reasons why OTK children were late enrolled into the first grade or why they had to repeat a school year are unknown. Based on the findings, it can only be speculated that OTK children may be delayed in their cognitive development which leads to a late enrollment in the first grade or to a repetition of a school year. This corresponds to a study by Urschitz et al. (2003). In a cohort of 1,144 German third graders, these authors reported that poor academic performance significantly increased with age for all school subjects (i.e., mathematics, science, reading, spelling, and handwriting) and that especially children who had repeated a school year were more likely to show poor academic performance (Urschitz et al., 2003). These findings are in line with other studies (Jaekel et al., 2015; Martin, 2009). Interestingly, in the cross-sectional study OTK girls outperformed OTK boys which corresponds to Urschitz et al. (2003). They observed that boys showed a larger prevalence for poor academic performance than girls, except for mathematics (Urschitz et al., 2003). This divergence between OTK girls and OTK boys might be attributed to biological maturation. As girls enter the adolescent growth spurt approximately two years earlier (i.e., at age 10) compared to boys (i.e., at age 12) (Malina et al., 2004), the better performance of OTK girls compared to OTK boys might be influenced by biological maturation.

Well in line with hypothesis 4, YTK children showed better performances than keyage-children especially in coordination. Again, the exact reasons why YTK children were early enrolled into the first grade or why they skipped a school year are unknown. Based on these findings, it can be speculated that YTK children may be accelerated in their cognitive development which leads to an early enrollment in the first grade or to a skipping of a school year. This assumption is reinforced by the finding that YTK children scored the best performance in the coordination task which requires a high level of cognitive functions. In addition, a study by Martin (2009) indicated a similar direction by revealing that in a cohort of 3,684 Australian high school students, YTK children showed better academic performances compared to keyage-children.

In summary, the cross-sectional study (Publication III) was the first study that analyzed differences in physical fitness development of YTK and OTK children compared to keyage-children. Schools, (physical education) teachers, politicians and decision makers, as well as parents should consider that OTK children often showed poorer performance in physical fitness than keyage-children. Therefore, it is necessary to particularly consider this "minority" to specifically provide health and fitness programs to reduce the gap in fitness experienced during prior years.

6.4 Limitations

Publications I, II, and III contributing to this cumulative doctoral thesis have a few potential limitations that warrant discussion. Only the main limitations of all three publications will be mentioned here. A detailed discussion of all limitations can be found within the respective sections of each publication.

In the systematic review (Publication I) biological maturity or anthropometric data (e.g., body mass, body height, or sitting height) were not included in the analysis because most of the considered studies did not report such data. There is strong evidence that physical fitness is associated with biological maturity as more mature adolescents showed better performance in physical fitness compared to less mature adolescents (Jones et al., 2000). In addition, a trend towards earlier maturation (i.e., age of puberty onset) was observed in previous centuries (Bellis et al., 2006; Herman-Giddens, 2006) which is accompanied by an increase in physical fitness which may have affected and altered secular trends. Furthermore, each reviewed study applied different study designs (e.g., instruments, tests, environmental conditions, methods etc.) for assessing physical fitness which resulted in inter-study variance, particularly in cardiorespiratory endurance and relative muscle strength. To better examine secular trends in physical fitness, future studies should apply the same test procedures to representative samples in children and adolescents from around the world (Fraser et al., 2019). In addition, all reviewed studies provided mean values of secular trends. However, trends in mean values might be systematically biased if there have been concurrent trends in skewness whereas median values will not be biased. Thus, future studies should also report measures of centrality (e.g., medians), variability (e.g., standard deviations), and asymmetry (e.g., skewness).

In the two cross-sectional studies (Publications II and III), the main limitations are of a methodological nature. From the cross-sectional study it remains unknown whether the linear age gains for tests held at the individual level or resulted from averaging over individual differences in non-linear growth curves (Publication II). Furthermore, several additional variables (e.g., physical activity, sedentary behavior, socioeconomic indicators) which influence physical fitness (Golle et al., 2014) were not included but could reveal an additional insight into the development of physical fitness (Publications II and III). Moreover, anthropometric data such as body mass, body height, and sitting height were not measured in the cross-sectional studies so that associations between physical fitness, anthropometric data, and biological maturation, could not be analyzed (Jones et al., 2000). These data would have provided additional insight as there is strong evidence that physical fitness is influenced by anthropometric data (Ceschia et al., 2016; Deforche et al., 2003; Sacchetti et al., 2012) and biological maturation (Malina et al., 2004) (Publications II and III). Lastly, performance of keyage-children in Publication III was predicted based on a linear extrapolation used in Publication II. However, it remains unknown if this linear extrapolation exactly fitted with the data of keyage-children aged 7.00 to 7.99 / 9.00 to 9.99 years as such cross-sectional data were missing (Publication III).

7. Conclusions

This doctoral thesis aggregated the findings of three publications that investigated the development of children's physical fitness over time as well as the effects of age, sex, and timing of school enrollment on the development of physical fitness with increasing age. The main results of the three publications can be summarized as follows:

1. There was a large initial improvement and an equally large subsequent decline between 1986 and 2010 as well as a stabilization between 2010 and 2015 in cardiorespiratory endurance, a general trend towards a small improvement in relative muscle strength, an overall small negative quadratic trend for proxies of muscle power, and a small-to-medium improvement in speed since 2002. Consequently, there was no generic secular trend across the different physical fitness components and years.
2. Even in a single prepubertal year of life physical fitness performance developed linearly with increasing chronological age. Boys showed better performances than girls in all physical fitness components. In addition, there was no evidence for an interaction of age and sex - with each other or with the components contrasts despite an abundance of statistical power.
3. The physical fitness components varied in the size of sex and age effects, with sex differences being larger for components requiring muscle mass and being smaller for components requiring motor coordination and with age differences being larger for components requiring motor coordination.
4. In the cross-sectional study performance in cardiorespiratory endurance declined and performance in speed increased between 2011 and 2019. In contrast, the initial increase in coordination, powerLOW, and powerUP was followed by a decrease in recent years.
5. OTK children showed poorer performance compared to keyage-children, especially in coordination and OTK girls showed better performances than OTK boys. In contrast, YTK children outperformed keyage-children, especially in coordination.

8. Practical implications

The present doctoral thesis and its findings aimed to fill the identified gaps in the literature regarding the development of children's physical fitness over time as well as the effects of age, sex, and timing of school enrollment on the development of physical fitness with increasing age. Several practical implications of high interest for schools, (physical education) teachers, politicians and decision makers, coaches, researchers, and parents can be drawn based on the presented results to prevent potential deficits in physical fitness as early as possible and consequently ensure a holistic development and a lifelong healthy life.

Due to the varying secular trends in physical fitness observed in Publications I and II as well as the literature, it is recommended to promote initiatives for physical activity and physical fitness for children and adolescents to prevent the adverse effects on health and well-being. More precisely, public health initiatives should specifically consider exercising cardiorespiratory endurance and muscle strength because cardiorespiratory endurance (Mintjens et al., 2018) and muscle strength (García-Hermoso et al., 2019; Rodrigues de Lima et al., 2020; Smith et al., 2014) showed strong positive associations with markers of health (i.e., BMI, fat mass, waist circumference, bone mass) in children and adolescents. Moreover, high levels of muscle strength are required for motor skill learning (Granacher et al., 2016). In this context, schools offer a suitable setting for the promotion of a physically active lifestyle (Marta et al., 2012) because this is the only setting that reaches all children regardless of their socioeconomic background (United Nations Educational Scientific and Cultural Organization, 2015). Thus, factors such as active education, active recess, or active transportation to school, etc. should be considered within schools to reduce sedentary behavior and promote physical activity (Kriemler et al., 2011; Meyer et al., 2014; Parrish et al., 2020). Moreover, parents and families should also be included in such activities (Faigenbaum, MacDonald, et al., 2020; Faigenbaum, Rebullido, et al., 2020; Faigenbaum & Myer, 2012; Palomäki et al., 2015).

In addition, the role of physical education lessons needs to be improved to further physical activity and physical fitness levels (Kriemler et al., 2011; Meyer et al., 2013; Meyer et al., 2014). Many studies showed the positive influence of resistance training (Drenowatz & Greier, 2018; Faigenbaum & MacDonald, 2017; Faigenbaum & Myer, 2010; Lesinski et al., 2020; Myers et al., 2017) and endurance training (Armstrong, 2016; Armstrong et al., 2011) on children's and adolescents' physical fitness. Based on these studies, muscle strengthening and aerobic exercises can be integrated in the curricula as research (Cox et al., 2020; Faigenbaum et al., 2015, 2019; Fu et al., 2019; Granacher et al., 2011; López-Gil et al., 2021; Loras, 2020;

Pozuelo-Carrascosa et al., 2018; Villa-González et al., 2022) indicated that well-structured and well-designed physical education classes positively influenced children's and adolescents' physical fitness. Furthermore, qualified physical education teachers are needed who know how to adequately structure motivating, joyful, and varied physical activity and physical fitness interventions for children and adolescents (Faigenbaum, MacDonald, et al., 2020). All these aspects imply that children and adolescents who regularly participate in physical activity (e.g., through physical education classes, in a sport club or free play) can acquire complex movements and sport skills (Faigenbaum & Bruno, 2017). Such interventions should ideally begin in primary schools as children at this stage of life are at an optimal age for acquiring motor skills (Brian et al., 2020; Faigenbaum & Myer, 2012).

Based on the findings from Publications I and II it is recommended to regularly analyze secular trends according to the specific physical fitness components and the monitored time span to detect changes in physical fitness as early as possible and consequently start interventions as early as possible to prevent, if necessary, adverse health outcomes. Schools offer a suitable setting for the evaluation of physical fitness because they reach all children during their development, growth, and formation of habits. Accordingly, findings from Publications II and III showed that the EMOTIKON test battery is easy-to-administer, cost effective, and requires only minimal school specific equipment (e.g., medicine ball, measuring tape, stopwatch, pylons). Therefore, physical education teachers, coaches, or researchers can use the EMOTIKON test battery to assess children's physical fitness and detect secular trends in children's physical fitness.

According to Ortega et al. (2011) physical fitness data of an individual should be compared with normative values of a sex and age-matched similar general population. However, such normative values are usually only available on annual or semi-annual basis. The findings of Publications II implied that physical education teachers, coaches, or researchers can utilize a proportional adjustment to individually interpret physical fitness of prepubertal school-aged children. Furthermore, Publication III revealed that schools, (physical education) teachers, politicians and decision makers, as well as parents should consider that OTK children often showed poorer performance in physical fitness than keyage-children. Thus, it is necessary to specifically consider this "minority" and provide additional health and fitness programs to reduce their deficits in physical fitness experienced during prior years to guarantee a holistic development. Moreover, raw data of Publication III can be used to calculate age-, sex- and timing of school enrollment-specific percentile values because reference values for the grading of physical fitness is only available for keyage-children. Those percentile values can be used by (physical education) teachers to individually interpret and grade children's physical fitness status.

9. Future directions

This doctoral thesis also offers a basis for future research projects in the field of children's development of physical fitness. Even though this thesis attempted to fill the recognized gaps in the literature regarding the development of children's physical fitness over time as well as the effects of age, sex, and timing of school enrollment on the development of physical fitness with increasing age, some important issues remain unresolved. Therefore, future studies should address the following issues:

1. Future studies on secular trends in physical fitness of children and adolescents are advised to examine the biological maturity status, if possible, to analyze the effect of biological maturity status on secular trends in physical fitness.
2. In order to better analyze secular trends in physical fitness, future studies should apply identical test procedures (i.e., tests, instruments, methods, environmental conditions etc.) among representative samples of children and adolescents from around the world to reduce inter-study variance.
3. Future studies on secular trends should consider middle- and low-income countries and country-specific representative data since existing studies on secular trends are mainly available for high-income countries such as the United States, Australia, and Europe and are usually not representative of the population of other countries.
4. Future studies on secular trends are advised to analyze measures of centrality (e.g., medians), variability (e.g., standard deviations), and asymmetry (e.g., skewness) since trends in mean values might be systematically biased.
5. Since there is only limited and inconsistent evidence for secular trends in the complex and multidimensional construct of coordination, future studies should especially focus on secular trends in coordination to further understand secular trends in coordination.
6. Future studies on children's development of physical fitness should consider several important variables (e.g., anthropometric measures [e.g., body mass, body height], status of biological maturity [e.g., peak height velocity, secondary sex characteristics], physical activity, sedentary behavior, socioeconomic indicators) which influence physical fitness to provide an additional insight into the development of physical fitness.

10. References

Albon, H. M., Hamlin, M. J., & Ross, J. J. (2010). Secular trends and distributional changes in health and fitness performance variables of 10-14-year-old children in New Zealand between 1991 and 2003. *British Journal of Sports Medicine, 44*(4), 263–269. https://doi.org/10.1136/bjsm.2008.047142

Albrecht, C., Hanssen-Doose, A., Bös, K., Schlenker, L., Schmidt, S., Wagner, M., Will, N., & Worth, A. (2016). Motorische Leistungsfähigkeit von Kindern und Jugendlichen in Deutschland: Eine 6-Jahres-Kohortenstudie im Rahmen des Motorik-Moduls (MoMo). *Sportwissenschaft*, *46*(4), 294–304. https://doi.org/10.1007/s12662-016-0421-4

Altman, D. G., & Bland, J. M. (1995). Absence of evidence is not evidence of absence. *Bmj*, *311*(19), 485. https://doi.org/10.1136/bmj.311.7003.485

Ardila, A., Rosselli, M., Matute, E., & Inozemtseva, O. (2011). Gender Differences in Cognitive Development. *Developmental Psychology, 47*(4), 984–990. https://doi.org/10.1037/a0023819

Armstrong, N., & Welsman, J. (2007). Aerobic fitness: What are we measuring? *Medicine and Sport Science*, *50*, 5–25. https://doi.org/10.1159/000101073

Armstrong, N. (2016). Aerobic Fitness and Training in Children and Adolescents. *Pediatric Exercise Science*, *28*(1), 7–10. https://doi.org/10.1123/pes.2016-0004

Armstrong, N., Tomkinson, G. R., & Ekelund, U. (2011). Aerobic fitness and its relationship to sport, exercise training and habitual physical activity during Youth. *British Journal of Sports Medicine*, *45*(11), 849–858. https://doi.org/10.1136/bjsports-2011-090200

Augste, C., & Künzell, S. (2014). Seasonal variations in physical fitness among elementary school children. *Journal of Sports Sciences*, *32*(5), 415–423. https://doi.org/10.1080/02640414.2013.830189

Bellis, M. A., Downing, J., & Ashton, J. R. (2006). Adults at 12? Trends in puberty and their public health consequences. *Journal of Epidemiology and Community Health*, *60*(11), 910–911. https://doi.org/10.1136/jech.2006.049379

Bennie, J. A., Faulkner, G., & Smith, J. J. (2021). The epidemiology of muscle-strengthening activity among adolescents from 28 European countries. *Scandinavian Journal of Public Health*, *50*(2), 295–302. https://doi.org/10.1177/14034948211031392

Bentham, J., Di Cesare, M., 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*(10113), 2627–2642. https://doi.org/10.1016/S0140-6736(17)32129-3

Beunen, G., & Thomis, M. (2000). Muscular strength development in children and adolescents. *Pediatric Exercise Science*, *12*(2), 174–197. https://doi.org/10.1123/pes.12.2.174

Bezanson, J., Edelman, A., Karpinski, S., & Shah, V. B. (2017). Julia: A fresh approach to numerical computing. *SIAM Review*, *59*(1), 65–98. https://doi.org/10.1137/141000671

Bös, K. (2009). *Deutscher Motorik-Test 6-18*. Feldhaus Edition Czwalina.

Brian, A., Getchell, N., True, L., De Meester, A., & Stodden, D. F. (2020). Reconceptualizing and Operationalizing Seefeldt's Proficiency Barrier: Applications and Future Directions. *Sports Medicine*, *50*(11), 1889–1900. https://doi.org/10.1007/s40279-020-01332-6

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. *British Journal of Sports Medicine*, *54*(24), 1451–1462. https://doi.org/10.1136/bjsports-2020-102955

Caspersen, C., Powell, K., & Christenson, G. (1985). Physical Activity, Exercise, and Physical Fitness: Definitions and Distinctions for Health-Related Research CARL. *Notes and Queries*, *100*(2), 125–131. https://doi.org/10.1093/nq/s9-IX.228.365-f

Castro-Piñero, J., González-Montesinos, J. L., Keating, X. D., Mora, J., Sjöström, M., & Ruiz, J. R. (2010). Percentile values for running sprint field tests in children ages 6-17 years: Influence of weight status. *Research Quarterly for Exercise and Sport*, *81*(2), 143–151. https://doi.org/10.1080/02701367.2010.10599661

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*(2), 98–108. https://doi.org/10.1136/bjsports-2011-090218

Ceschia, A., Giacomini, S., Santarossa, S., Rugo, M., Salvadego, D., Da Ponte, A., Driussi, C., Mihaleje, M., Poser, S., & Lazzer, S. (2016). Deleterious effects of obesity on physical fitness in pre-pubertal children. *European Journal of Sport Science*, *16*(2), 271–278. https://doi.org/10.1080/17461391.2015.1030454

Cobley, S., Baker, J., Wattie, N., & McKenna, J. (2009). Annual age-grouping and athlete development: A meta-analytical review of relative age effects in sport. *Sports Medicine*, *39*(3), 235–256. https://doi.org/10.2165/00007256-200939030-00005

Costa, A. M., Costa, M. J., Reis, A. A., Ferreira, S., Martins, J., & Pereira, A. (2017). Secular Trends in Anthropometrics and Physical Fitness of Young Portuguese School-Aged Children. *Acta Médica Portuguesa*, *30*(2), 108–114. https://doi.org/10.20344/amp.7712

Cox, A., Fairclough, S. J., Kosteli, M. C., & Noonan, R. J. (2020). Efficacy of School-Based Interventions for Improving Muscular Fitness Outcomes in Adolescent Boys: A Systematic Review and Meta-analysis. *Sports Medicine*, *50*(3), 543–560. https://doi.org/10.1007/s40279-019-01215-5

Cureton, K. J., Sparling, P. B., Evans, P. W., Johnson, S. M., Kong, U. D., & Purvis, J. W. (1978). Effect of experimental alterations in excess weight on aerobic capacity and distance running performance. *Medicine and Science in Sports*, *10*(3), 194–199.

De Miguel-Etayo, P., 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*, 57–66. https://doi.org/10.1038/ijo.2014.136

Deforche, B., Lefevre, J., De Bourdeaudhuij, I., Hills, A. P., Duquet, W., & Bouckaert, J. (2003). Physical fitness and physical activity in obese and nonobese Flemish youth. *Obesity Research*, *11*(3), 434–441. https://doi.org/10.1038/oby.2003.59

Dobosz, J., Mayorga-Vega, D., & Viciana, J. (2015). Percentile values of physical fitness levels among polish children aged 7 to 19 years – a population-based study. *Central European Journal of Public Health*, *23*(4), 340–351. https://doi.org/10.21101/cejph.a4153

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. *Sports Medicine*, *50*(6), 1129–1144. https://doi.org/10.1007/s40279-020-01265-0

Drenowatz, C., & Greier, K. (2018). Resistance Training in Youth - Benefits and Characteristics. *Journal of Biomedicine*, *3*, 32–39. https://doi.org/10.7150/jbm.25035

Drenowatz, C., Hinterkörner, F., & Greier, K. (2021). Physical Fitness and motor competence in upper Austrian elementary school children — study protocol and preliminary findings of a state-wide fitness testing program. *Frontiers in Sports and Active Living, 3,* 1–11. https://doi.org/10.3389/fspor.2021.635478

Ducher, G., Bass, S. L., Naughton, G. A., Eser, P., Telford, R. D., & Daly, R. M. (2009). Overweight children have a greater proportion of fat mass relative to muscle mass in the upper limbs than in the lower limbs: Implications for bone strength at the distal forearm. *American Journal of Clinical Nutrition*, *90*(4), 1104–1111. https://doi.org/10.3945/ajcn.2009.28025

Eberhardt, T., Niessner, C., Oriwol, D., Buchal, L., Worth, A., & Bös, K. (2020). Secular trends in physical fitness of children and adolescents: A review of large-scale epidemiological studies published after 2006. *International Journal of Environmental Research and Public Health*, *17*(16), 1–17. https://doi.org/10.3390/ijerph17165671

Faigenbaum, A. D., & Bruno, L. E. (2017). A Fundamental Approach for Treating Pediatric Dynapenia in Kids. *ACSM's Health and Fitness Journal*, *21*(4), 18–24. https://doi.org/10.1249/FIT.0000000000000312

Faigenbaum, A. D., Bush, J. A., McLoone, R. P., Kreckel, M. C., Farrell, A., Ratamess, N. A., & Kang, J. (2015). Benefits of strength and skill-based training during primary school physical education. *Journal of Strength and Conditioning Research*, *29*(5), 1255–1262. https://doi.org/10.1519/JSC.0000000000000812.

Faigenbaum, A. D., & MacDonald, J. P. (2017). Dynapenia: it's not just for grown-ups anymore. *Acta Paediatrica, International Journal of Paediatrics*, *106*(5), 696–697. https://doi.org/10.1111/apa.13797

Faigenbaum, A. D., MacDonald, J. P., Carvalho, C., & Rebullido, T. R. (2020). The Pediatric Inactivity Triad: A Triple Jeopardy for Modern Day Youth. *ACSM's Health and Fitness Journal*, *24*(4), 10–17. https://doi.org/10.1249/FIT.0000000000000584

Faigenbaum, A. D., & Myer, G. D. (2010). Pediatric resistance training: Benefits, concerns, and program design considerations. *Current Sports Medicine Reports*, *9*(3), 161–168. https://doi.org/10.1249/JSR.0b013e3181de1214

Faigenbaum, A. D., & Myer, G. D. (2012). Exercise deficit disorder in youth: Play now or pay later. *Current Sports Medicine Reports*, *11*(4), 196–200. https://doi.org/10.1249/JSR.0b013e31825da961.

Faigenbaum, A. D., Rebullido, T. R., & Chulvi-Medrano, I. (2020). Youth Physical Activity Is All About the "F-Words." *Strength & Conditioning Journal*, *42*(6), 2–6. https://doi.org/10.1519/ssc.0000000000000530

Faigenbaum, A. D., Rebullido, T. R., & MacDonald, J. P. (2018). Pediatric Inactivity Triad: A Risky PIT. *Current Sports Medicine Reports*, *17*(2), 45–47. https://doi.org/10.1249/JSR.0000000000000450

Faigenbaum, A. D., Rebullido, T. R., Peña, J., & Chulvi-Medrano, I. (2019). Resistance Exercise for the Prevention and Treatment of Pediatric Dynapenia. *Journal of Science in Sport and Exercise*, *1*(3), 208–216. https://doi.org/10.1007/s42978-019-00038-0

Fernandez-Santos, J. R., Ruiz, J. R., Cohen, D. D., Gonzalez-Montesinos, J. L., & Castro-Pinero, J. (2015). Reliability and validity of tests to assess lower-body muscular power in children. *Journal of Strength and Conditioning Research*, *29*(8), 2277–2285. https://doi.org/10.1519/JSC.0000000000000864

Finger, J. D., Varnaccia, G., Borrmann, A., Lange, C., & Mensink, G. B. M. (2018). Körperliche Aktivität von Kindern und Jugendlichen in Deutschland -Querschnittergebnisse aus KiGGS Welle 2 und Trends. *Journal of Health Monitoring*, *3*(1), 24–31. https://doi.org/10.17886/RKI-GBE-2018-006

Fraser, B. J., Blizzard, L., Buscot, M.-J., Schmidt, M. D., Dwyer, T., Venn, A. J., & Magnussen, C. G. (2021). Muscular strength across the life course: The tracking and trajectory patterns of muscular strength between childhood and mid-adulthood in an Australian cohort. *Journal of Science and Medicine in Sport*, *24*(7), 696–701. https://doi.org/10.1016/j.jsams.2021.01.011

Fraser, B. J., Blizzard, L., Tomkinson, G. R., Lycett, K., Wake, M., Burgner, D., Ranganathan, S., Juonala, M., Dwyer, T., Venn, A. J., Olds, T., & Magnussen, C. G. (2019). The great leap backward: changes in the jumping performance of Australian children aged 11−12-years between 1985 and 2015. *Journal of Sports Sciences*, *37*(7), 748–754. https://doi.org/10.1080/02640414.2018.1523672

Fraser, B. J., Schmidt, M. D., Huynh, Q. L., Dwyer, T., Venn, A. J., & Magnussen, C. G. (2017). Tracking of muscular strength and power from youth to young adulthood: Longitudinal findings from the Childhood Determinants of Adult Health Study. *Journal of Science and Medicine in Sport*, *20*(10), 927–931. https://doi.org/10.1016/j.jsams.2017.03.021

Fu, Y., Burns, R. D., Constantino, N., Fitzsimmons, J., & Zhang, P. (2019). Effect of the

Resistance Exercise on Elementary School Students' Physical Fitness. *Journal of Science in Sport and Exercise*, *1*(2), 184–191. https://doi.org/10.1007/s42978-019-0022-7

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*, 17566. 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. *Scientific Reports*, *12*, 7801. https://doi.org/10.1038/s41598-022-11710-x

Fühner, T., Kliegl, R., Arntz, F., Kriemler, S., & Granacher, U. (2021). An Update on Secular Trends in Physical Fitness of Children and Adolescents from 1972 to 2015: A Systematic Review. *Sports Medicine*, *51*(2), 303–320. https://doi.org/10.1007/s40279-020-01373-x

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. *Sports Medicine*, *49*(7), 1079–1094. https://doi.org/10.1007/s40279-019-01098-6

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*(1), 1–8. 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*(11), e0142393. https://doi.org/10.1371/journal.pone.0142393

Granacher, U., Lesinski, M., Büsch, D., Muehlbauer, T., Prieske, O., Puta, C., Gollhofer, A., & Behm, D. G. (2016). Effects of resistance training in youth athletes on muscular fitness and athletic performance: A conceptual model for long-term athlete development. *Frontiers in Physiology*, *7*, 164. https://doi.org/10.3389/fphys.2016.00164

Granacher, U., Muehlbauer, T., Doerflinger, B., Strohmeier, R., & Gollhofer, A. (2011). Promoting Strength and Balance in Adolescents during physical education: effects of s short-term resistance training. *Journal of Strength and Conditioning Research*, *25*(4), 940–949. https://doi.org/10.1519/JSC.0b013e3181c7bb1e

Guthold, R., Stevens, G. A., Riley, L. M., & Bull, F. C. (2020). Global trends in insufficient physical activity among adolescents: a pooled analysis of 298 population-based surveys with 1·6 million participants. *The Lancet Child and Adolescent Health*, *4*(1), 23–35.

https://doi.org/10.1016/S2352-4642(19)30323-2

Hallal, P. C., Andersen, L. B., Bull, F. C., Guthold, R., Haskell, W., Ekelund, U., Alkandari, J. R., Bauman, A. E., Blair, S. N., Brownson, R. C., Craig, C. L., Goenka, S., Heath, G. W., Inoue, S., Kahlmeier, S., Katzmarzyk, P. T., Kohl, H. W., Lambert, E. V., Lee, I. M., … Wells, J. C. (2012). Global physical activity levels: Surveillance progress, pitfalls, and prospects. *The Lancet*, *380*(9838), 247–257. https://doi.org/10.1016/S0140-6736(12)60646-1

Herman-Giddens, M. E. (2006). Recent data on pubertal milestones in United States children: The secular trend toward earlier development. *International Journal of Andrology*, *29*(1), 241–246. https://doi.org/10.1111/j.1365-2605.2005.00575.x

Hjorth, M. F., Chaput, J. P., Michaelsen, K., Astrup, A., Tetens, I., & Sjödin, A. (2013). Seasonal variation in objectively measured physical activity, sedentary time, cardio-respiratory fitness and sleep duration among 8-11 year-old Danish children: A repeated-measures study. *BMC Public Health*, *13*(1), 1–10. https://doi.org/10.1186/1471-2458-13-808

Jaekel, J., Strauss, V. Y. C., Johnson, S., Gilmore, C., & Wolke, D. (2015). Delayed school entry and academic performance: A natural experiment. *Developmental Medicine and Child Neurology*, *57*(7), 652–659. https://doi.org/10.1111/dmcn.12713

Jones, M. A., Hitchen, P. J., & Stratton, G. (2000). The importance of considering biological maturity when assessing physical fitness measures in girls and boys aged 10 to 16 years. *Annals of Human Biology*, *27*(1), 57–65. https://doi.org/10.1080/030144600282389

Kalman, M., Inchley, J., Sigmundova, D., Iannotti, R. J., Tynjälä, J. A., Hamrik, Z., Haug, E., & Bucksch, J. (2015). Secular trends in moderate-to-vigorous physical activity in 32 countries from 2002 to 2010: A cross-national perspective. *European Journal of Public Health*, *25*, 37–40. https://doi.org/10.1093/eurpub/ckv024

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*(1), 150–156. https://doi.org/10.1007/BF00964130

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*, *38*(16), 1913–1923. https://doi.org/10.1080/02640414.2020.1763764

Katzmarzyk, P. T., Malina, R. M., Song, T. M. K., & Bouchard, C. (1998). Physical activity and health-related fitness in youth: a multivariate analysis. *Medicine and Science in Sports and Exercise, 30*(5), 709–714. https://doi.org/10.1097/00005768-199805000-00011

Koutsandréou, F., Wegner, M., Niemann, C., & Budde, H. (2016). Effects of motor versus cardiovascular exercise training on children's working memory. *Medicine and Science in Sports and Exercise*, *48*(6), 1144–1152. https://doi.org/10.1249/MSS.0000000000000869

Kriemler, S., Meyer, U., Martin, E., Van Sluijs, E. M. F., 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. *British Journal of Sports Medicine*, *45*(11), 923–930. https://doi.org/10.1136/bjsports-2011-090186

Larsen, L. R., Kristensen, P. L., Junge, T., Rexen, C. T., & Wedderkopp, N. (2015). Motor Performance as Predictor of Physical Activity in Children: The CHAMPS Study-DK. *Medicine and Science in Sports and Exercise*, *47*(9), 1849–1856. https://doi.org/10.1249/MSS.0000000000000604

Laurson, K. R., Baptista, F., Mahar, M. T., Welk, G. J., & Janz, K. F. (2022). Long Jump, Vertical Jump, and Vertical Jump Power Reference Curves for 10-18 Year Olds. *Measurement in Physical Education and Exercise Science*, 1–9. https://doi.org/10.1080/1091367x.2021.2017291

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*(2), 93–101. https://doi.org/10.1080/02640418808729800

Lesinski, M., Herz, M., Schmelcher, A., & Granacher, U. (2020). Effects of Resistance Training on Physical Fitness in Healthy Children and Adolescents: An Umbrella Review. *Sports Medicine*, *50*(11), 1901–1928. https://doi.org/10.1007/s40279-020-01327-3

Liberati, A., Altman, D. G., Tetzlaff, J., Mulrow, C., Gøtzsche, P. C., Ioannidis, J. P. A., Clarke, M., Devereaux, P. J., Kleijnen, J., & Moher, D. (2009). The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: Explanation and elaboration. *PLoS Medicine*, *6*(7), e1000100. https://doi.org/10.1371/journal.pmed.1000100

Lisowski, P., Kantanista, A., & Bronikowski, M. (2020). Are there any differences between first grade boys and girls in physical fitness, physical activity, BMI, and sedentary behavior? Results of HCSC study. *International Journal of Environmental Research and Public Health*, *17*(3), 1109. https://doi.org/10.3390/ijerph17031109

López-Gil, J. F., Cavero-Redondo, I., Tárraga-López, P. J., de Camargo, E. M., Sequí-Domínguez, I., Yuste Lucas, J. L., Renato Cavichiolli, F., & García-Hermoso, A. (2021). Intensity of physical activity in physical education classes and school recesses and its associations with body mass index and global fitness score in spanish schoolchildren. *Applied Sciences (Switzerland)*, *11*(23), 11337. https://doi.org/10.3390/app112311337

Loras, H. (2020). The Effects of Physical Education on Motor Competence in Children and Adolescents: a systematic review and meta-analysis. *Sports*, *88*(8), 1–14. https://doi.org/10.3390/sports8060088

Ludyga, S., Gerber, M., Pühse, U., Looser, V. N., & Kamijo, K. (2020). Systematic review and meta-analysis investigating moderators of long-term effects of exercise on cognition in healthy individuals. *Nature Human Behaviour*, *4*(6), 603–612. https://doi.org/10.1038/s41562-020-0851-8

Malina, R. M., Bouchard, C., & Bar-Or, O. (2004). *Growth, Maturation, and Physical Activity* (2nd ed.). Human Kinetics.

Marshall, W. A., & Tanner, J. M. (1969). Variations in pattern of pubertal changes in girls. *Archives of Disease in Childhood*, *44*(235), 291–303. https://doi.org/http://dx.doi.org/10.1136/adc.44.235.291

Marshall, W. A., & Tanner, J. M. (1970). Variations in pattern of pubertal changes in boys. *Archives of Disease in Childhood*, *45*(239), 13–23. https://doi.org/http://dx.doi.org/10.1136/adc.45.239.13

Marta, C. C., Marinho, D. A., Barbosa, T. M., Izquierdo, M., & Marques, M. C. (2012). Physical Fitness Differences Between Prepubescent Boys and Girls. *Journal of Strength and Conditioning Research*, *26*(7), 1756–1766. https://doi.org/10.1519/JSC.0b013e31825bb4aa

Martin, A. J. (2009). Age Appropriateness and Motivation, Engagement, and Performance in High School: Effects of Age Within Cohort, Grade Retention, and Delayed School Entry. *Journal of Educational Psychology*, *101*(1), 101–114. https://doi.org/10.1037/a0013100

Mayorga-Vega, D., Casado-Robles, C., Viciana, J., & López-Fernández, I. (2019). Daily step-based recommendations related to moderate-to-vigorous physical activity and sedentary behavior in adolescents. *Journal of Sports Science and Medicine*, *18*(4), 586–595.

Meyer, U., Roth, R., Zahner, L., Gerber, M., Puder, J. J., Hebestreit, H., & Kriemler, S. (2013). Contribution of physical education to overall physical activity. *Scandinavian Journal of Medicine and Science in Sports*, *23*(5), 600–606. https://doi.org/10.1111/j.1600-0838.2011.01425.x

Meyer, U., Schindler, C., Zahner, L., Ernst, D., Hebestreit, H., Van Mechelen, W., Rocca, H. P. B. La, Probst-Hensch, N., Puder, J. J., & Kriemler, S. (2014). Long-term effect of a school-based physical activity program (KISS) on fitness and adiposity in children: A cluster-randomized controlled trial. *PLoS ONE*, *9*(2), e87929. https://doi.org/10.1371/journal.pone.0087929

Ministerium für Bildung Jugend und Sport. (2018). *Gesetz über die Schulen im Land Brandenburg*. https://bravors.brandenburg.de/gesetze/bbgschulg

Ministerium für Bildung Jugend und Sport. (2021a). *Einschulungen nach Schulform, Art der Einschulung und Schuljahr*. https://mbjs.brandenburg.de/media_fast/bb2.a.5978.de/z_einsch2_2020_21.pdf

Ministerium für Bildung Jugend und Sport. (2021b). *Wiederholer (Nichtversetzte oder freiwillige Wiederholer) nach Schulstufe, Jahrgangsstufe, Schulform und Schuljahr*. https://mbjs.brandenburg.de/media_fast/bb2.a.5978.de/z_wied_P_2020_21.pdf

Mintjens, S., Menting, M. D., Daams, J. G., van Poppel, M. N. M., Roseboom, T. J., & Gemke, R. J. B. J. (2018). Cardiorespiratory Fitness in Childhood and Adolescence Affects Future Cardiovascular Risk Factors: A Systematic Review of Longitudinal Studies. *Sports Medicine*, *48*(11), 2577–2605. https://doi.org/10.1007/s40279-018-0974-5

Moher, D., Liberati, A., Tetzlaff, J., Altman, D. G., Altman, D., Antes, G., Atkins, D., Barbour, V., Barrowman, N., Berlin, J. A., Clark, J., Clarke, M., Cook, D., D'Amico, R., Deeks, J. J., Devereaux, P. J., Dickersin, K., Egger, M., Ernst, E., … Tugwell, P. (2009). Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *PLoS Medicine*, *6*(7), e1000097. https://doi.org/10.1371/journal.pmed.1000097

Musch, J., & Grondin, S. (2001). Unequal Competition as an Impediment to Personal Development: A Review of the Relative Age Effect in Sport. *Developmental Review*, *21*, 147–167. https://doi.org/https://doi.org/10.1006/drev.2000.0516

Myers, A. M., Beam, N. W., & Fakhoury, J. D. (2017). Resistance training for children and adolescents. *Translational Pediatrics*, *6*(3), 137–143. https://doi.org/10.21037/tp.2017.04.01

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*, 1–9. https://doi.org/10.3389/fpubh.2020.00458

Olds, T., & Dollman, J. (2004). Are Changes in Distance-Run Performance of Australian Children Between 1985 and 1997 Explained by Changes in Fatness? *Pediatric Exercise Science*, *46*(15), 201–209. https://doi.org/10.1111/j.1746-1561.1991.tb06021.x

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*(1), 20–29. https://doi.org/10.1136/bjsm.2009.062679

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

Overman, W. H. (2004). Sex differences in early childhood, adolescence, and adulthood on cognitive tasks that rely on orbital prefrontal cortex. *Brain and Cognition*, *55*(1), 134–147. https://doi.org/10.1016/S0278-2626(03)00279-3

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*(6), 588–596. https://doi.org/10.1080/02640414.2014.951874

Parrish, A. M., Chong, K. H., Moriarty, A. L., Batterham, M., & Ridgers, N. D. (2020). Interventions to Change School Recess Activity Levels in Children and Adolescents: A Systematic Review and Meta-Analysis. *Sports Medicine*, *50*(12), 2145–2173. https://doi.org/10.1007/s40279-020-01347-z

Pate, R. R. (1991). Health-related measures of children ’ s physical fitness. *Journal of School Health*, *61*(5), 231–233. https://doi.org/10.1111/j.1746-1561.1991.tb06021.x

Poitras, V. J., Gray, C. E., Borghese, M. M., Carson, V., Chaput, J.-P., Janssen, I., Katzmarzyk, P. T., Pate, R. R., Connor Gorber, S., Kho, M. E., Sampson, M., & Tremblay, M. S. (2016). Systematic review of the relationships between objectively measured physical activity and health indicators in school-aged children and youth. *Applied Physiology, Nutrition, and Metabolism*, *41*(6), 197–239. https://doi.org/10.1139/apnm-2015-0663

Pozuelo-Carrascosa, D. P., García-Hermoso, A., Álvarez-Bueno, C., Sánchez-López, M., & Martinez-Vizcaino, V. (2018). Effectiveness of school-based physical activity programmes on cardiorespiratory fitness in children: A meta-Analysis of randomised controlled trials. *British Journal of Sports Medicine*, *52*(19), 1234–1240. https://doi.org/10.1136/bjsports-2017-097600

R Core Team. (2020). *R: A language and environment for statistical computing*. Foundation for Statistical Computing, Vienna, Austria. http://www.r-project.org/

Ramos-Sepúlveda, J. A., Ramírez-Vélez, R., Correa-Bautista, J. E., Izquierdo, M., & García-Hermoso, A. (2016). Physical fitness and anthropometric normative values among Colombian-Indian schoolchildren. *BMC Public Health*, *16*, 1–15. https://doi.org/10.1186/s12889-016-3652-2

Rodrigues de Lima, T., Custódio Martins, P., Henrique Guerra, P., & Augusto Santos Silva, D. (2020). Muscular Fitness and Cardiovascular Risk Factors in Children and Adolescents: A Systematic Review. *Journal of Strength and Conditioning Research*, *34*(8), 2394–2406. https://doi.org/10.1519/JSC.0000000000002840

Roriz de Oliveira, M. S., Seabra, A., Freitas, D., Eisenmann, J. C., & Maia, J. (2014). Physical fitness percentile charts for children aged 6-10 from Portugal. *Journal of Sports Medicine and Physical Fitness*, *54*, 780–792.

Rosenfield, R. L., B.Lipton, R., & Drum, M. L. (2009). Thelarche, pubarche, and menarche attainment in children with normal and elevated body mass index. *Pediatrics*, *123*(1), 84–88. https://doi.org/10.1542/peds.2008-0146

Round, J. M., Jones, D. A., Honour, J. W., & Nevill, A. M. (1999). Hormonal factors in the development of differences in strength between boys and girls during adolescence: A longitudinal study. *Annals of Human Biology*, *26*(1), 49–62. https://doi.org/10.1080/030144699282976

Rütten, A., & Pfeifer, K. (2016). *Nationale Empfehlungen für Bewegung und Bewegungsförderung*. FAU.

Sacchetti, R., Ceciliani, A., Garulli, A., Masotti, A., Poletti, G., Beltrami, P., & Leoni, E. (2012). Physical fitness of primary school children in relation to overweight prevalence and physical activity habits. *Journal of Sports Sciences*, *30*(7), 633–640.

https://doi.org/10.1080/02640414.2012.661070

Sandercock, G. R. H., Taylor, M., Voss, C., Qgunleye, A. A., Cohen, D. D., & Parry, D. A. (2013). Quantification of the relative age effect in three indices of physical performance. *Journal of Strength and Conditioning Research*, *27*(12), 3293–3299. https://doi.org/10.1519/JSC.0b013e318291b28d

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*(16), 1510–1518. https://doi.org/10.1080/02640414.2014.906046

Schmidt, M., Egger, F., Benzing, V., Jäger, K., Conzelmann, A., Roebers, C. M., & Pesce, C. (2017). Disentangling the relationship between children's motor ability, executive function and academic achievement. *PLoS ONE*, *12*(8), e0182845. https://doi.org/10.1371/journal.pone.0182845

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

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. *Sports Medicine*, *44*(9), 1209–1223. https://doi.org/10.1007/s40279-014-0196-4

Stratton, G., & Oliver, J. L. (2020). The impact of growth and maturation on physical performance. In R. S. Lloyd & J. L. Oliver (Eds.), *Strength and conditioning for young athletes: science and application* (2nd ed., pp. 3–20). Routledge.

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*(6), 736–746. https://doi.org/10.1080/17461391.2015.1088577

Telama, R., Yang, X., Leskinen, E., Kankaanpää, A., Hirvensalo, M., Tammelin, T., Viikari, J. S. A., & Raitakari, O. T. (2014). Tracking of physical activity from early childhood through youth into adulthood. *Medicine and Science in Sports and Exercise*, *46*(5), 955–962. https://doi.org/10.1249/MSS.0000000000000181

Tomkinson, G. R., & Olds, T. (2007). Secular changes in aerobic fitness test performance of

Australasian children and adolescents. *Medicine and Sport Science*, *50*(314), 168–182. https://doi.org/10.1159/000101361

Tomkinson, G. R. (2007). Global changes in anaerobic fitness test performance of children and adolescents (1958-2003). *Scandinavian Journal of Medicine and Science in Sports*, *17*(5), 497–507. https://doi.org/10.1111/j.1600-0838.2006.00569.x

Tomkinson, G. R., & Olds, T. S. (2007). Secular changes in pediatric aerobic fitness test performance: the global picture. In *Pediatric fitness. Secular trends and geographic variability* (Vol. 50, pp. 46–66). Karger. https://doi.org/10.1159/000101358

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*(22), 1445–1456. https://doi.org/10.1136/bjsports-2017-098253

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*, *53*(8), 478–486. https://doi.org/10.1136/bjsports-2017-097982

Tomkinson, G. R, Kaster, T., Dooley, F. L., Fitzgerald, J. S., Annandale, M., & Ferrar, K. (2021). Temporal Trends in the Standing Broad Jump Performance of 10, 940,801 Children and Adolescents Between 1960 and 2017. *Sports Medicine*, *51*(3), 531–548. https://doi.org/10.1007/s40279-020-01394-6

United Nations Educational Scientific and Cultural Organization. (2015). *International Charter of Physical Education and Sport*. United Nations Educational, Scientific and Cultural Organization.

Urschitz, M. S., Guenther, A., Eggebrecht, E., Wolff, J., Urschitz-Duprat, P. M., Schlaud, M., & Poets, C. F. (2003). Snoring, intermittent hypoxia and academic performance in primary school children. *American Journal of Respiratory and Critical Care Medicine*, *168*(4), 464–468. https://doi.org/10.1164/rccm.200212-1397OC

Venckunas, T., Emeljanovas, A., Mieziene, B., & Volbekiene, V. (2017). Secular trends in physical fitness and body size in Lithuanian children and adolescents between 1992 and 2012. *Journal of Epidemiology and Community Health*, *71*(2), 181–187. https://doi.org/10.1136/jech-2016-207307

Villa-González, E., Barranco-Ruiz, Y., García-Hermoso, A., & Faigenbaum, A. D. (2022). Efficacy of school-based interventions for improving muscular fitness outcomes in children: A systematic review and meta-analysis. *European Journal of Sport Science*, 1–34. https://doi.org/10.1080/17461391.2022.2029578

Viru, A., Loko, J., Harro, M., Volver, A., Laaneots, L., & Viru, M. (1999). Critical Periods in the Development of Performance Capacity During Childhood and Adolescence. *European Journal of Physical Education*, *4*(1), 75–119. https://doi.org/10.1080/1740898990040106

Viru, A., Loko, J., Volver, A., Laaneots, L., Karelson, K., & Viru. (1998). Age periods of accelerated improvemnt of muscle strength, power, speed and endurance in the age interval 6-18 yeras. *Biology of Sport*, *15*(4), 211–227.

Woll, A., Kurth, B. M., Opper, E., Worth, A., & Bös, K. (2011). The "Motorik-Modul" (MoMo): Physical fitness and physical activity in German children and adolescents. *European Journal of Pediatrics*, *170*(9), 1129–1142. https://doi.org/10.1007/s00431-010-1391-4

World Health Organization. (2010). *Global Recommendations on Physical Activity for Health*. World Health Organization.

World Health Organization. (2018). *Global action plan on physical activity 2018–2030: more active people for a healthier world.* World Health Organization.

Wrotniak, B. H., Epstein, L. H., Dorn, J. M., Jones, K. E., & Kondilis, V. A. (2006). The relationship between motor proficiency and physical activity in children. *Pediatrics, 118*(6), e1758–e1765. https://doi.org/10.1542/peds.2006-0742

Acknowledgements

My special thanks go to my supervisor Prof. Dr. Urs Granacher for his tireless support and helpfulness as well as his excellent supervision of my dissertation. Thank you, Mr. Granacher, for numerous constructive and motivating discussions. Your immense knowledge of training and movement science has always opened up new perspectives for me.

In addition, I would like to express my sincere thanks to my second supervisor Prof. Dr. Reinhold Kliegl for his valuable statistical support and hours of statistical training. Thank you, Reinhold, for your perseverance in always bringing *R* closer to me.

Furthermore, I would like to thank the Ministry of Education, Youth and Sports, especially Eckhard Drewicke and Martina Schünemann, for their great interest in EMOTIKON and for always wanting to advance the project. Moreover, I would like to thank all primary schools in the Federal State of Brandenburg, Germany for participating in EMOTIKON, because only through the annual data collection this thesis was made possible.

Further thanks go to my dear (former) colleagues Kathleen, Eva, Melanie, Fabian, Arnd, Markus, Tom, Lukas, Adrian, Helmi, Chrissi, Sophia, Willi, Steffen, Robert, Kim, and Paula. Whenever I had any questions - no matter in which direction - you were always there. You have always enriched the working day with countless canteen meals, coffee rounds, and a humorous environment.

In addition, I thank Eva, Paula, and Markus who took their valuable time to proofread this thesis.

A big thank you also goes to my dearest friends Miri and Michaela. You were always there for me and provided the necessary distraction from my thesis. Thank you for always having an open ear for me and for motivating me!

My biggest thanks go to my parents Annette and Dieter Fühner and my partner Alexander Schlößer! Dear Mom and Dad: thank you for giving me the opportunity to pursue my academic career and for always supporting me in many ways. Without you, the scientific path I have chosen would not have been possible! Dear Alex: thank you for always being there for me, for sharing the joy in my successes and for always building me up and believing in me when I face any hurdle!

Authors‘ contribution

The present doctoral thesis is designed as a cumulative dissertation. Consequently, three scientific articles have been submitted to international peer-reviewed journals and accepted for publication. According to the doctoral degree regulations of the Human Sciences Faculty of the University of Potsdam (§ 7 (4), sentence No. 2), significant contributions to the articles from the respective co-authors were acknowledged and confirmed by each co-author.

Erklärung des Promovenden / der Promovendin
Zum eigenen Anteil an den vorgelegten wissenschaftlichen Abhandlungen
mit zwei oder mehr Autor(inn)en
(kumulative Dissertation)

Name der Promovendin:
Thea Heidi Fühner

Titel der Dissertation:
Secular trends, age, sex, and timing of school enrollment effects on physical fitness in children and adolescents

Betreuer:
Prof. Dr. Urs Granacher

Wissenschaftliche Abhandlung 1

Titel	An update on secular trends in physical fitness of children and adolescents from 1972 to 2015: a systematic review
Autor(ein)	Fühner, T., Kliegl, R., Arntz, F., Kriemler, S., & Granacher, U.
Journal	Sports Medicine
	Publikationsstatus (bitte ankreuzen)
( )	Nicht eingereicht
( )	Eingereicht
( )	In Begutachtung
( )	Angenommen
(x)	Veröffentlicht / Publikationsjahr: 2021

Beschreibung des eigenen Anteils, wenn keine Alleinautorenschaft vorliegt:

Zur Entwicklung dieses Artikels habe ich in folgender Art und Weise beigetragen:

- Mitwirkung an der Konzeption der genauen Fragestellung und des Studiendesigns
- Klärung der Umsetzung und des statistischen Modells
- Eigenständige Durchführung der Datenaufnahme/Untersuchung
- Datenaufbereitung und -analyse
- Dateninterpretation
- Erstellung des Manuskriptes sowie Abstimmung und Einarbeitung der Überarbeitungsvorschläge
- Revision der Gutachterkommentare

[redacted]

Reinhold Kliegl Fabian Arntz Susi Kriemler Urs Granacher

Erklärung des Promovenden / der Promovendin
Zum eigenen Anteil an den vorgelegten wissenschaftlichen Abhandlungen
mit zwei oder mehr Autor(inn)en
(kumulative Dissertation)

Name der Promovendin:
Thea Heidi Fühner

Titel der Dissertation:
Secular trends, age, sex, and timing of school enrollment effects on physical fitness in children and adolescents

Betreuer:
Prof. Dr. Urs Granacher

Wissenschaftliche Abhandlung 2

Titel	Age and sex effects in physical fitness components of 108,295 third graders including 515 primary schools and 9 cohorts
Autor(ein)	Führer, T., Granacher, U., Golle, K., & Kliegl, R.
Journal	Scientific Reports
	Publikationsstatus (bitte ankreuzen)
( )	Nicht eingereicht
( )	Eingereicht
( )	In Begutachtung
( )	Angenommen
(x)	Veröffentlicht / Publikationsjahr: 2021

Beschreibung des eigenen Anteils, wenn **keine** Alleinautorenschaft vorliegt:

Zur Entwicklung dieses Artikels habe ich in folgender Art und Weise beigetragen:

- Mitwirkung an der Konzeption der genauen Fragestellung und des Studiendesigns
- Klärung der Umsetzung und des statistischen Modells
- Eigenständige Durchführung der Datenaufnahme/Untersuchung
- Datenaufbereitung und -analyse
- Dateninterpretation
- Erstellung des Manuskriptes sowie Abstimmung und Einarbeitung der Überarbeitungsvorschläge
- Revision der Gutachterkommentare

[redacted]

Urs Granacher | Kathleen Golle | Reinhold Kliegl

Erklärung des Promovenden / der Promovendin
Zum eigenen Anteil an den vorgelegten wissenschaftlichen Abhandlungen
mit zwei oder mehr Autor(inn)en
(kumulative Dissertation)

Name der Promovendin:
Thea Heidi Fühner

Titel der Dissertation:
Secular trends, age, sex, and timing of school enrollment effects on physical fitness in children and adolescents

Betreuer:
Prof. Dr. Urs Granacher

Wissenschaftliche Abhandlung 3

Titel	Effect of timing of school enrollment on physical fitness in third graders
Autor(ein)	Fühner, T., Granacher, U., Golle, K., & Kliegl, R.
Journal	Scientific Reports
	Publikationsstatus (bitte ankreuzen)
( )	Nicht eingereicht
( )	Eingereicht
( )	In Begutachtung
( )	Angenommen
(x)	Veröffentlicht / Publikationsjahr: 2022

Beschreibung des eigenen Anteils, wenn **keine** Alleinautorenschaft vorliegt:

Zur Entwicklung dieses Artikels habe ich in folgender Art und Weise beigetragen:

- Mitwirkung an der Konzeption der genauen Fragestellung und des Studiendesigns
- Klärung der Umsetzung und des statistischen Modells
- Eigenständige Durchführung der Datenaufnahme/Untersuchung
- Datenaufbereitung und -analyse
- Dateninterpretation
- Erstellung des Manuskriptes sowie Abstimmung und Einarbeitung der Überarbeitungsvorschläge
- Revision der Gutachterkommentare

[redacted]

Urs Granacher | Kathleen Golle | Reinhold Kliegl

Publication I

AN UPDATE ON SECULAR TRENDS IN PHYSICAL FITNESS OF CHILDREN AND ADOLESCENTS FROM 1972 TO 2015: A SYSTEMATIV REVIEW

Thea Fühner[1], Reinhold Kliegl[1], Fabian Arntz[1], Susi Kriemler[2], Urs Granacher[1]

[1] University of Potsdam, Research Focus Cognition Sciences, Division of Training and Movement Science, Potsdam, Germany

[2] University of Zurich, Epidemiology, Biostatistics and Prevention Institute, Zurich, Switzerland

Reference

Fühner, T., Kliegl, R., Arntz, F., Kriemler, S., & Granacher, U. (2021a). An update on secular trends in physical fitness of children and adolescents from 1972 to 2015: a systematic review. *Sports Medicine, 51*, 303-320. https://doi.org/10.1007/s40279-020-01373-x

The final publication is available at:

https://link.springer.com/article/10.1007/s40279-020-01373-x

An Update on Secular Trends in Physical Fitness of Children and Adolescents from 1972 to 2015:

A Systematic Review

Running title: Secular Trends in Physical Fitness of Children and Adolescents

Thea Fühner[1]

Reinhold Kliegl[1]

Fabian Arntz[1]

Susi Kriemler[2]

Urs Granacher[1]

[1]Division of Training and Movement Sciences, Research Focus Cognition Sciences, University of Potsdam, Am Neuen Palais 10, Building 12, 14469 Potsdam, Germany

[2]Epidemiology, Biostatistics and Prevention Institute, University of Zurich, Hirschengraben 84, 8001 Zurich, Switzerland

Corresponding author and contact details

Prof. Urs Granacher, PhD

Division of Training and Movement Sciences

Research Focus Cognition Sciences

University of Potsdam

Am Neuen Palais 10, Building 12

14469 Potsdam

Germany

Tel: +49-331-977 1543

E-mail address: urs.granacher@uni-potsdam.de

ORCID ID: orcid.org/0000-0002-7095-813X

Abstract

Background There is evidence that physical fitness of children and adolescents (particularly cardiorespiratory endurance) has declined globally over the past decades. Ever since the first reports on negative trends in physical fitness, efforts were undertaken by for instance the World Health Organization (WHO) to promote physical activity and fitness in children and adolescents. Therefore, it is timely to re-analyze the literature to examine whether previous reports on secular declines in physical fitness are still detectable or whether they need to be updated.

Objectives The objective of this systematic review is to provide an 'update' on secular trends in selected components of physical fitness (i.e., cardiorespiratory endurance, relative muscle strength, proxies of muscle power, speed) in children and adolescents aged 6–18 years.

Data Sources A systematic computerized literature search was conducted in the electronic databases PubMed and Web of Science to locate studies that explicitly reported secular trends in physical fitness of children and adolescents.

Study Eligibility Criteria Studies were included in this systematic review if they examined secular trends between at least two time points across a minimum of five years. In addition, they had to document secular trends in any measure of cardiorespiratory endurance, relative muscle strength, proxies of muscle power or speed in apparently healthy children and adolescents aged 6–18 years.

Study Appraisal and Synthesis Methods The included studies were coded for the following criteria: nation, physical fitness component (cardiorespiratory endurance, relative muscle strength, proxies of muscle power, speed), chronological age, sex (boys vs. girls), and year of assessment. Scores were standardized (i.e., converted to z-scores) with sample-weighted means and standard deviations, pooled across sex and year of assessment within cells defined by study, test, and children´s age.

Results The original search identified 524 hits. In the end, 22 studies met the inclusion criteria for review. The observation period was between 1972 and 2015. Fifteen of the 22 studies used tests for cardiorespiratory endurance, eight for relative muscle strength, eleven for proxies of muscle power, and eight for speed. Measures of cardiorespiratory endurance exhibited a large initial increase and an equally large subsequent decrease, but the decrease appears to have reached a floor for all children between 2010 and 2015. Measures of relative muscle strength showed a general trend towards a small increase. Measures of proxies of muscle

power indicated an overall small negative quadratic trend. For measures of speed, a small-to-medium increase was observed in recent years.

Limitations Biological maturity was not considered in the analysis because biological maturity was not reported in most included studies.

Conclusions Negative secular trends were particularly found for cardiorespiratory endurance between 1986 and 2010–12, irrespective of sex. Relative muscle strength and speed showed small increases while proxies of muscle power declined. Although the negative trend in cardiorespiratory endurance appears to have reached a floor in recent years, because of its association with markers of health, we recommend further initiatives in PA and fitness promotion for children and adolescents. More specifically, public health efforts should focus on exercising cardiorespiratory endurance to prevent adverse health effects (i.e., overweight and obesity) and muscle strength to lay a foundation for motor skill learning.

Key Points

This systematic review documents a large initial increase and an equally large subsequent decrease for cardiorespiratory endurance, but the decrease appears to have reached a floor for all children between 2010 and 2015. Relative muscle strength showed a small increase. Proxies of muscle power indicated an overall small negative quadratic trend. Speed showed small-to-medium increases in recent years.

Because of the different trends in physical fitness, we recommend that already existing programs with the goal to promote PA and fitness should be maintained. More specifically, public health efforts should focus on exercising cardiorespiratory endurance to prevent adverse health effects (overweight, obesity) and muscle strength to lay a foundation for motor skill learning.

1. Introduction

The World Health Organization (WHO) recommends at least 60 min of moderate-to-vigorous physical activity (PA) daily and additionally muscle and bone strengthening activities three times per week for children and adolescents aged 5-17 years [1]. Recently published studies showed that a majority of children and adolescents (~80%) around the globe do not meet the recommended level of 60 minutes PA per day [2–4]. Notwithstanding, if children and adolescents do not adhere to WHO recommendations [1], they are supposed to suffer from ‘exercise deficit disorder’ including all negative health consequences [5].

Childhood is an important developmental stage to acquire fundamental movement skills through daily PA in order to obtain motor skill competence and movement confidence. Children who do not gain such competencies due to sedentariness are more likely to experience adverse health outcomes later in life [5]. Furthermore, it has been postulated that a physically active lifestyle during childhood and adolescence is robust and tracks into adulthood [6–9]. For instance, Telama et al. [9] conducted a 27-year follow-up measurement of 3,596 Finnish boys and girls aged 3–18 years and reported that PA behavior develops during childhood and tracks into adulthood with moderate-to-high stability (stability coefficients ≥ 0.60).

There is evidence from cross-sectional [10] and longitudinal studies [11] of an association between levels of PA and physical fitness. According to Caspersen et al. [12], physical fitness can be categorized into health- (e.g., cardiorespiratory endurance, muscle strength, etc.) and skill-related (e.g., speed, power, etc.) components of physical fitness. Wrotniak et al. [10] used accelerometers to objectively measure PA and observed that time in moderate-to-vigorous PA positively correlated with measures of muscular strength (standing broad jump; $r = 0.40$) and speed (running speed; $r = -0.36$) in children aged 8–10 years. These results were confirmed in a three-year longitudinal study in which positive associations were reported between time spent in moderate-to-vigorous PA measured through accelerometers and different components of physical fitness (i.e., muscle strength (handgrip, $\beta = 0.06$)), proxies of muscle power (vertical jump, $\beta = 0.04$)) in children aged 6–12 years [11].

Evidence-based research indicates that physical fitness is a powerful marker of health in children and adolescents [6]. Especially, cardiorespiratory endurance [13] and muscle strength [8] have been found to be positively associated with markers of health in children and adolescents. In a systematic review, Mintjens et al. [13] reported that performance levels in cardiorespiratory endurance were positively related with body mass index, waist circumference, body fatness, and prevalence of metabolic syndrome. Furthermore, Garzia-Hermoso et al. [8]

reported positive associations between measures of muscle strength and body mass index, skinfold thickness, insulin resistance, triglycerides, cardiovascular disease risk score, and bone mineral density. Testing of physical fitness is easy-to-administer, reliable and valid which is why it should be extensively implemented to receive information on performance development and health of children and adolescents [14].

There is already evidence available in the literature on secular declines in physical fitness of children and adolescents. Tomkinson et al. [15, 16] summarized the existing literature on this topic for the timespan between 1958 and 2003. They included physical fitness data of 25,000,000–50,000,000 children and adolescents aged 6–19 years living in 27 countries across five geographical regions. Over the entire analysis period (1958–2003), cardiorespiratory endurance declined by -0.36% per annum (p.a.). A more in-depth analysis indicated that between 1958 and 1970, cardiorespiratory endurance improved by +0.61% p.a. Thereafter, performance declined by -0.54% p.a. Findings for proxies of muscle power and speed were different in as much as there was an overall positive trend for both qualities (proxies of muscle power: +0.03% p.a.; speed +0.04% p.a.) between 1958 and 2003. When looking at the period from 1958 to 1985, improvements were found for proxies of muscle power (+0.44% p.a.) and speed (+0.27% p.a.). These were followed by annual declines in proxies of muscle power (-0.20% p.a.) and speed (-0.08% p.a.). Recently, Tomkinson et al. [17] published an update on secular trends in cardiorespiratory endurance restricting the analysis to data on the 20-m shuttle run test that were published between 1981 and 2014. In their meta-analysis, Tomkinson et al. [17] summarized data of almost 1,000,000 children and adolescents aged 9–17 years living in 19 different countries. According to their results, the international rate in cardiorespiratory endurance decline has slowed and stabilized since the turn of the century [17]. Furthermore, a recently published systematic analysis of secular trends on handgrip strength including data of 2,000,000 children and adolescents aged 9 to 17 years living in 19 countries showed that the international rate of improvement in handgrip strength progressively increased between 1967 and 2017 [18].

In summary, there is evidence that physical fitness of children and adolescents (particularly cardiorespiratory endurance) has declined globally, particularly in western industrialized countries until the turn of the century and stabilized ever since. These trend analyses in physical fitness were the starting point to specifically promote PA and physical fitness in children and adolescents, especially by the WHO [19].

Given the effort of the WHO over the past years to promote PA and physical fitness in children and adolescents [19], it appears important to regularly update these secular trend analyses. Considering that those updates are available only for cardiorespiratory endurance using estimated V̇O2peak from the 20-m shuttle run test [17] and absolute muscle strength data from handgrip dynamometry [18], it is timely to re-analyze the literature and examine secular trends for other components of physical fitness such as relative muscle strength, proxies of muscle power, and speed. It is also important to broaden the perspective on cardiorespiratory endurance, i.e., to include also tests other than the 20-m shuttle run test in the analyses.

The aim of this systematic review was to provide an 'update' on secular trends in selected components of physical fitness (i.e., cardiorespiratory endurance, relative muscle strength, proxies of muscle power, speed) in children and adolescents aged 6–18 years. With reference to recent studies [17, 18] and due to ongoing public health efforts to promote PA and fitness [19], we hypothesized a positive trend in physical fitness development of children and adolescents over time. We additionally expected that the reported secular trends are heterogeneous both, in direction and magnitude and specific to the respective physical fitness component under consideration. This could be due to a multitude of tests that were used to assess the different components of physical fitness.

2. Methods

This systematic review was carried out in accordance with the Preferred Reporting Items for Systematic Reviews and Meta-Analysis (PRISMA) statement guidelines [20, 21].

2.1 Literature Search

The authors conducted a systematic computerized literature search in the electronic databases PubMed and Web of Science to locate studies that explicitly reported secular trends in physical fitness of apparently healthy children and adolescents. The literature search period covered publications until April 2019. An updated search in July 2020 could not identify any additional hits. The following Boolean search strategy was used to identify studies related to secular trends in physical fitness of children and adolescents. Proximity operators ("*") were applied to search for root words: ("physical fitness" OR "cardiorespiratory endurance" OR "muscular endurance" OR "muscular strength" OR "body composition" OR flexibility OR agility OR balance OR coordination OR speed OR power OR "reaction time") AND (child OR children OR youth OR adolescent OR adolescents OR adolescence) AND ("secular change*" OR "secular trend*" OR "secular decline*" OR "temporal trend*" OR "temporal change*" OR "temporal decline*").

In addition, the following filters were activated for PubMed: species: humans; ages: birth–18 years (children), 13–18 years (adolescents). The applied search syntax for PubMed was adapted for the Web of Science database so that the abbreviation ”TS=” (for Topic) was placed in front of each bracket. Furthermore, reference lists of each article as well as relevant review articles and meta-analyses [15, 16, 22–31] were cross-referenced/screened to identify suitable adequate references to be included in this systematic review (see flow chart Figure 1). We additionally scrutinized the database ‘Cochrane Library‘, but could not identify any additional hits.

2.2 Selection Criteria

Studies were integrated in this systematic review if they provided relevant information with regards to a modified PICOS approach [21]. The considered parameters were population, outcomes, and study design. The following predefined inclusion criteria were selected: (1) population: apparently healthy children and adolescents with a mean age of 6–18 years; differentiated age ranges not spanning more than three years (e.g., 12- to 14-year olds) according to Tomkinson, Olds and Borms [32]; a sample distinguished by sex allowing sex-specific analysis as recommended by Tomkinson [16]; (2) intervention: not possible; (3) comparator: not possible; (4) outcomes: secular trends in selected physical fitness components (i.e., cardiorespiratory endurance, relative muscle strength, proxies of muscle power, speed) expressed as mean, standard deviation / standard error, and sample size according to these studies [15, 16, 23]; (5) study design: in accordance with Jürimäe et al. [33], studies which spanned at least two time points across a minimum of five years; only published articles/datasets with full-text availability; articles published in English or German. Studies were excluded for the following reasons: (1) population: patients or outside the mean age range from 6 to 18 years; large age ranges spanning more than three years (e.g., 12- to 15-year olds) [32]; children with an elite athletic background; children who live in specific areas (e.g., rural areas, Inuit) since their data are typically not representative for the general population [22]; a sample not distinguished by sex as recommended by Tomkinson [16]; (2) intervention: not possible; (3) comparator: not possible; (4) outcomes: assessment of secular trends not related to physical fitness; authors reported only aggregated physical fitness scores; missing data concerning means, standard deviation / standard error or sample size; (5) study design: samples which did not span at least two time points across a minimum of five years [33]; no full-text availability; articles published in languages other than English or German; data that were reported in several publications [32].

Two independent reviewers (TF, FA) screened potentially relevant papers by analyzing titles, abstracts, and full texts to determine study eligibility. If TF and FA did not reach agreement, UG was consulted for clarification.

2.3 Coding of Studies

All included studies were coded for the following variables listed in Table 1. The author TF extracted the relevant data in an Excel spreadsheet from all included studies.

Table 1 Study Coding

Nation	Australia, Brazil, Canada, Czech Republic, Denmark, England, Finland, Belgium (Flanders), Germany, Hungary, Lithuania, Netherlands, New Zealand, Norway, Portugal, Sweden, USA
Physical fitness components	Cardiorespiratory endurance, relative muscle strength, proxies of muscle power, speed
Chronological age	6-18 years
Sex	Boys & girls
Trends	Positive, negative, stagnation

If studies included several tests of the same physical fitness component, only one test was used according to the following ranking (Table 2). The list was created through expert rating based on the factors ‘practical relevance’ and ‘frequency of usage’ of the respective test.

Table 2 Classification of the ranked physical fitness tests used in the included studies

Cardiorespiratory endurance	Relative muscle strength	Proxies of muscle power	Speed
6-min run*	Leg lift test*	Standing broad jump*	20-m sprint*
9-min run*	Sit-up[a*]	Sargent jump and reach test*	30-m sprint*
20-m shuttle run*	Push-up[b*]		40-m sprint*
Cooper test (12 min)*	Bent arm hang*		45,7-m (50-yd) sprint*
1200-m run*	Pull-up[c *]		50-m sprint*
1600-m run*	Arm-pull		4 x 9-m shuttle run
Maximal cycle ergometer*	Bench-press[d]		10 x 5-m shuttle run*
PWC 170 cycle ergometer*	Two-hand lift		
Submaximal cycle ergometer*			

[a] = maximum repetition; number at a rate of 25 lifts per minute or counts/30 sec, [b] = counts/40 sec, [c] = maximum repetition, [d] = number at a rate of 25 lifts per minute, * tests were included in analyses

If relevant data (e.g., means) were only reported in figures or graphs, the program GetData-Graph-Digitizer (http://www.getdata-graph-digitizer.com/index.php) was used for data extraction purposes.

2.4 Statistical Analyses

2.4.1 Computation of *z-scores* for Components of Physical Fitness

Physical fitness was assessed through the fitness components cardiorespiratory endurance, relative muscle strength, proxies of muscle power, and speed. For measures of relative muscle strength, proxies of muscle power, and cardiorespiratory endurance, high scores indicate good and for measures of speed high scores indicate poor physical fitness. To facilitate readability, speed outcomes were multiplied by -1. Thus, performance improvements over time were positive for all included outcome measures. The total sample size was calculated from the maximal sample sizes at age–time points–sex level for each included study.

As a common metric for tests, we converted means to z-scores based on sample-weighted means and standard deviations for cells defined by study, test, and age of children. Specifically, means and standard deviations were pooled across sex and year within these cells. For three studies, standard deviations were computed from standard errors [34–36]; for two studies, the number of participants for subgroups was extrapolated from aggregate numbers [37, 38]. Of note, Spengler et al. [39] provided means and standard deviations for their study in response to a personal request. For pre- and post-processing of results, custom R code and (mainly) the following R packages were used: *tidyverse* [40], *remef* [41], *sjPlot* [42], *cowplot* [43], and *broom.mixed* [44].

2.4.2 Statistical Inference with an Integrated Linear Mixed Model

We specified a linear mixed model (LMM) to estimate effects of sex and linear, quadratic, and cubic secular trends as nested under the four components of physical fitness (i.e., 8 x 4 = 32 fixed effects). For each observation, year of assessment was centred at 1990, i.e., at the midpoint of first-year assessments between 1974 and 2006 across studies. The nested model specification implies that for each of the four fitness components, we obtained an estimate for the year 1990 (i.e., equivalent to an intercept), an estimate for the main effect of sex, estimates for the secular trend (three parameters for linear, quadratic, and cubic component), and estimates for the interactions of sex and secular trend.

We tested the effects of sex and secular trends within each of the four physical fitness components because it was clear from the outset that the qualitative differences between the physical fitness components would yield trivial higher-order interactions. The limitation of the nested-model specification is that we do not obtain outcomes that report differences between these physical fitness components (e.g., sex x component or secular trend x component interactions). We considered separate LMMs for each physical fitness component. This approach, however, would not take dependencies between measures from different physical fitness components and between dependencies across years within studies into account. Moreover, the number of observations available for each LMM would substantially reduce statistical power for each of them.

The LMM was estimated with the *lmer()* function of the R-based *lme4* package [45, 46]. Model selection involved the determination of a random-effect structure that is supported by the data using random-effect principal component analysis [47, 48]. The final model included study-related variance components for cardiorespiratory endurance, relative muscle strength, sex, and linear yearly trends for cardiorespiratory endurance, relative muscle strength, and proxies of muscle power. The specification of correlation parameters for these variance components led to overparameterized models or did not significantly contribute to the goodness of fit as assessed with likelihood ratio tests. With this random-effect structure, none of the high-order interactions between sex and quadratic or cubic secular trends were significant.[1] Therefore, we removed these eight fixed effects from the LMM. The protocol of model selection is available as an Electronic Supplementary Material Appendix S2.

The LMM treated the 22 studies as levels of a random factor; they contributed a total of 652 observations. The LMM estimated 24 fixed effects, six variance components, and the residual variance (i.e., a total of 31 model parameters). A |z|-value > 2.0 (i.e., alpha of 5%) was used as criterion for rejection of the null hypothesis.

In response to reviewer requests, we specified two additional LMMs. First, we included age of children (dichotomized at 12.5 years) as an additional factor and documented this analysis in an Electronic Supplementary Material Appendix S1. The second request was to address possible geographical/cultural differences between studies from different countries. This analysis is also reported in an Electronic Supplementary Material Appendix S1.

3. Results

3.1 Study Characteristics

A total of 524 potentially relevant studies were identified in the electronic databases PubMed and Web of Science. Fourteen articles were found through other sources such as reference lists of relevant reviews articles or meta-analyses. After screening for titles, abstracts, and full texts, 22 studies were finally eligible for inclusion in this systematic review article. Figure 1 illustrates the respective flow chart.

[1] Fixed effects were not considered and not even inspected during selection of the random-effects structure supported by the data.

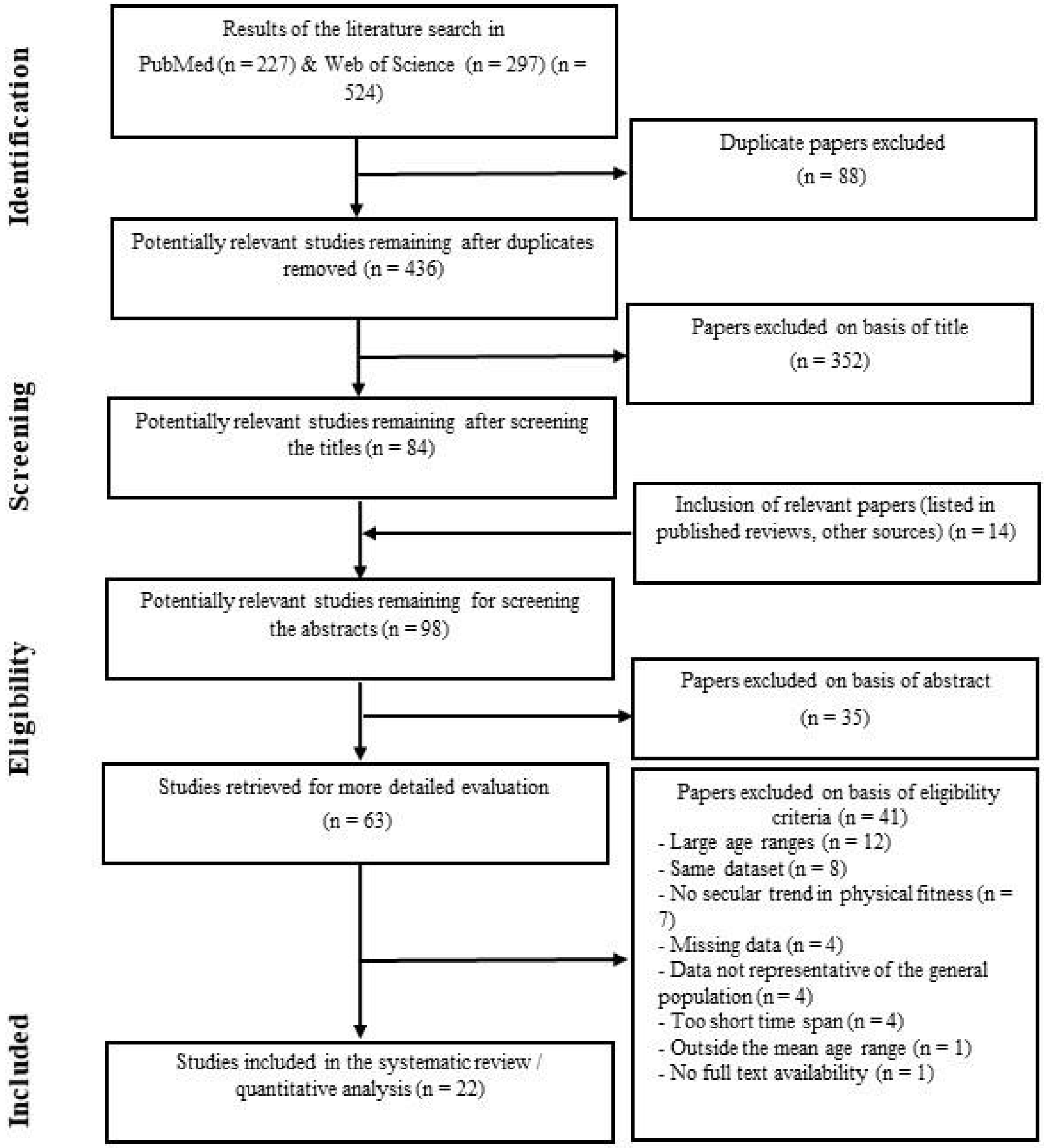

Fig. 1 Flow chart illustrating the search and selection process of this systematic review

Table 3 summarizes the characteristics of the included studies. Secular trends in physical fitness were analyzed for N = 96,522 children and adolescents aged 6-18 years living in 17 different countries mainly in high-income countries such as the United States, Australia, and Europe. Sample sizes ranged between 41 and 2,153 participants. Sixteen of the 22 studies were carried out in Europe [34, 37–39, 49–60], two in Australia [35, 61], one in the United States [36], one in Canada [62], one in New Zealand [63], and one in Brazil [64]. The time span varied from six to 35 years with a mean value of 20 years and a median of 21 years. Six studies reported data for several time points [34, 39, 49, 57, 58, 64].

To assess secular trends in physical fitness, 15 out of 22 studies used tests for cardiorespiratory endurance [37–39, 50, 51, 55–64], 11 used tests for proxies of muscle power [34–36, 50, 52–54, 56, 57, 60, 61], eight for relative muscle strength [34, 36, 39, 49, 50, 54, 57, 60], and eight for speed [34, 36, 39, 49, 51, 54, 57, 61]. A total of 26 different physical fitness tests were found in the included studies (Table 2).

Studies which assessed cardiorespiratory endurance included submaximal [56, 60, 63, 64] or maximal [55] tests on a cycle ergometer, the 20-m shuttle run test [57–59, 62] or time-related (6-min run test [39], 9-min run test [50], Cooper test [37, 38]) as well as distance-related cardiorespiratory endurance tests (1200-m run test [51], 1600-m run test [61]). Studies which assessed relative muscle strength comprised five different lower/upper limbs strength tests as well as tests for the assessment of trunk muscle strength. Accordingly, tests included the leg lift test [54], sit-up test [34, 50, 57], push-up test [39, 60], bent arm hang test [49], and pull-up test [36]. Studies which assessed proxies of lower limbs muscle power included either the standing broad jump test [34–36, 52, 53, 57, 60, 61] or the Sargent jump and reach test [50, 54, 56]. Studies which assessed speed comprised either linear sprint tests over short distances such as the 20-m sprint test [39], 30-m sprint test [51], the 40-m sprint test [34], the 45-m sprint test [36], the 50-m sprint test [61] or short shuttle runs such as the 10 x 5-m shuttle run test [49, 54, 57] (Table 3).

Table 3 Summary of the studies used in this systematic review. Studies were sorted chronologically according to the first year of measurement. The table illustrates the country of origin, the time points (years), the observational period (years), sex, the respective age range of the participating children and adolescents, the range of sample sizes, and the test(s) for which secular trends were reported

Reference	Country	Time points (years) Observational period in brackets (years)	Sex	Age range (years)	Range of sample sizes	Physical fitness component	Test(s)
Matton et al. [49]	Belgium (Flanders)	1972[a]/1980[b], 2005 (33[a]/25[b])	M, F	12 17[c]	161–2,153	Strength, speed	BAH, SHR10 x 5
Westerstahl et al. [50]	Sweden	1974, 1995 (21)	M, F	16	185–230	Strength, power, endurance	9 min, THL, SJAR, SU, BP
Reiff et al. [36]	USA	1975, 1985 (10)	M, F	10–17[c]	196–786	Strength, power, speed	PU, SHR4 x 9, 45,7 m, SBJ
Mészáros et al. [51]	Hungary	1975, 2000 (25)	M	10–13[c]	160–191	Speed, endurance	30 m, 1200 m
Krombholz [52]	Germany	1977, 2000 (23)	M, F	6–7[c]	80–220	Power	SBJ
Sedlak et al. [53]	Czech Republic	1977, 2012 (35)	M, F	6	133–178	Power	SBJ
Sziva et al. [38]	Hungary	1979, 2004 (25)	M	7–11[c]	152–158[d]	Endurance	12 min
De Moraes Ferrari et al. [64]	Brazil	1979, 1989, 1999, 2009 (30)	M, F	10–11	43–184	Endurance	SCE
Runhaar et al. [54]	Netherlands	1980, 2006 (26)	M, F	9–12[c]	41–505	Strength, power, speed	BAH, SHR10 x 5, LLT, AP, SJAR
Reed et al. [62]	Canada	1981, 2004 (23)	M, F	9 –11[c]	252-2,151[d]	Endurance	SHR 20 m
Dollmann et al. [61]	Australia	1985, 1997 (12)	M, F	10–11[c]	277–499	Power, speed, endurance	1600 m, 50 m, SBJ
Hardy et al. [35]	Australia	1985, 2015 (30)	M, F	9–15[c]	109–546	Power	SBJ
Wedderkopp et al. [55]	Denmark	1986, 1998 (12)	M, F	9	279–670	Endurance	MCE
McAnally et al. [63]	New Zealand	1987, 2011 (24)	M, F	15	157–436	Endurance	SCE
Aaberge & Mamen [56]	Norway	1988, 2001 (13)	M, F	15	77–106	Power, endurance	SCE, SJAR
Pampakas et al. [37]	Hungary	1989, 2004 (15)	M	7–10,5[c]	136–147[d]	Endurance	12 min
Venckunas et al. [57]	Lithuania	1992, 2002, 2012 (20)	M, F	11–18[c]	58–598	Strength, power, speed, endurance	SBJ, SHR10 x 5, BAH, SU, SHR 20 m
Costa et al. [34]	Portugal	1996, 2001, 2006, 2011 (15)	M, F	10–11	229–262	Strength, power, speed	SU, SBJ, 40 m
Sandercock et al. [58]	England	1998, 2008, 2014 (16)	M, F	10–11	150–158	Endurance	SHR 20 m
Palomäki et al. [59]	Finland	2003, 2010 (7)	M, F	15–16	640–1,142	Endurance	SHR 20 m
Albrecht et al. [60]	Germany	2005, 2011 (6)	M, F	11–13	352–466	Strength, power, endurance	SBJ, PHU, PWC
Spengler et al. [39]	Germany	2006-2015 (9)	M, F	6–7	220–274	Strength, speed, endurance	6 min, PHU, 20 m

M = male, F = female, BAH = bent arm hang, THL = two hand lift, SU = sit-up, BP = bench-press, PU = pull-up, PHU = push-up, LLT = leg lift test, AP = arm-pull, SBJ = standing broad jump, SJAR = Sargent Jump and Reach, 20 m = 20-m sprint, 30 m = 30-m sprint, 40 m = 40-m sprint, 45,7 m = 45,7-m (50-yd) sprint, 50 m = 50-m sprint, SHR4 x 9 = 4 x 9-m shuttle run, SHR10 x 5 = 10 x 5-m shuttle run, 6 min = 6-min run, 9 min = 9-min run, 12 min = 12-min run (Cooper Test), 1200 m = 1200-m run, 1600 m = 1600-m run, SHR 20 m = 20-m shuttle run, PWC = PWC 170 cycle ergometer, SCE = submaximal cycle ergometer, MCE = maximal cycle ergometer, [a] = for male, [b] = for female, [c] = data are available for each yearly ages, [d] = authors only reported the hole sample size

3.2 Secular Trends and Effects of Sex for Components of Physical Fitness

The 24 fixed-effect LMM estimates relating to sex, secular trend, and their interactions for each of the four components are displayed in Table 4 along with standard error, z-statistic, and p-value. In general, we observed the expected higher scores for boys than girls in all panels with effects ranging between 0.094 (relative muscle strength; not significant with $p < 0.07$) and 0.44 (cardiorespiratory endurance, $p < 0.001$). Interpretation of polynomial trends, however, must start with the highest-order significant trend or its interaction with sex; lower-order trends and main effects are usually qualified by them. In the following, we report statistics for the critical terms (lifted from Table 4); statistics for all lower-order terms are reported in the table.

Table 4 Fixed-effects estimates of linear mixed model

Component	Fixed-effects estimates	Standard error	z-values	p-values
Cardiorespiratory endurance	0.32595	0.12234	2.66	**0.008**
Sex	0.44032	0.05348	8.23	**<0.001**
Year (linear)	-0.02419	0.01027	-2.36	**0.018**
Year (quadratic)	-0.00365	0.00084	-4.35	**<0.001**
Year (cubic)	0.00012	0.00003	3.94	**<0.001**
Sex x year (linear)	-0.00796	0.00243	-3.28	**0.001**
Relative muscle strength	-0.17252	0.07174	-2.40	**0.016**
Sex	0.09356	0.05087	1.84	0.066
Year (linear)	0.01095	0.00720	1.52	0.128
Year (quadratic)	0.00067	0.00033	2.069	**0.039**
Year (cubic)	-0.00004	0.00002	-2.14	**0.032**
Sex x year (linear)	0.00066	0.00184	0.36	0.721
Proxies of muscle power	0.09430	0.03376	2.79	**0.005**
Sex	0.36078	0.04894	7.37	**<0.001**
Year (linear)	-0.00093	0.00536	-0.17	0.863
Year (quadratic)	-0.00060	0.00030	-2.04	**0.042**
Year (cubic)	0.00001	0.00002	0.70	0.486
Sex x year (linear)	0.00474	0.00169	2.80	**0.005**
Speed	0.03498	0.03732	0.94	0.349
Sex	0.18428	0.05015	3.67	**<0.001**
Year (linear)	-0.01822	0.00372	-4.90	**<0.001**
Year (quadratic)	-0.00044	0.00021	-2.11	**0.035**
Year (cubic)	0.00006	0.00001	4.60	**<0.001**
Sex x year (linear)	0.00277	0.00180	1.54	0.125

The coefficients define the functions that generate the secular trends shown in Fig.2a. Significant effects are in bold

Data and model parameters (up to the highest-order significant term) were used to generate the partial-effect functions for secular trends shown in Figure 2a. Partial effects no longer contain effects due to differences between studies in cardiorespiratory endurance, relative muscle strength, sex, or due to linear yearly trends for cardiorespiratory endurance, relative muscle strength, and proxies of muscle power. Points in the figure represent the observed sex x year means that result from averaging z-scores across observations within age x study/test cells. Note that corresponding means predicted by data and model parameters (i.e., the values underlying the function fit) would be closer to the function due to LMM shrinkage (i.e., adjustments for variability, number of observations, and deviation from the Grand Mean). Nevertheless, observed means are also in good correspondence with partial-effect Z-functions.

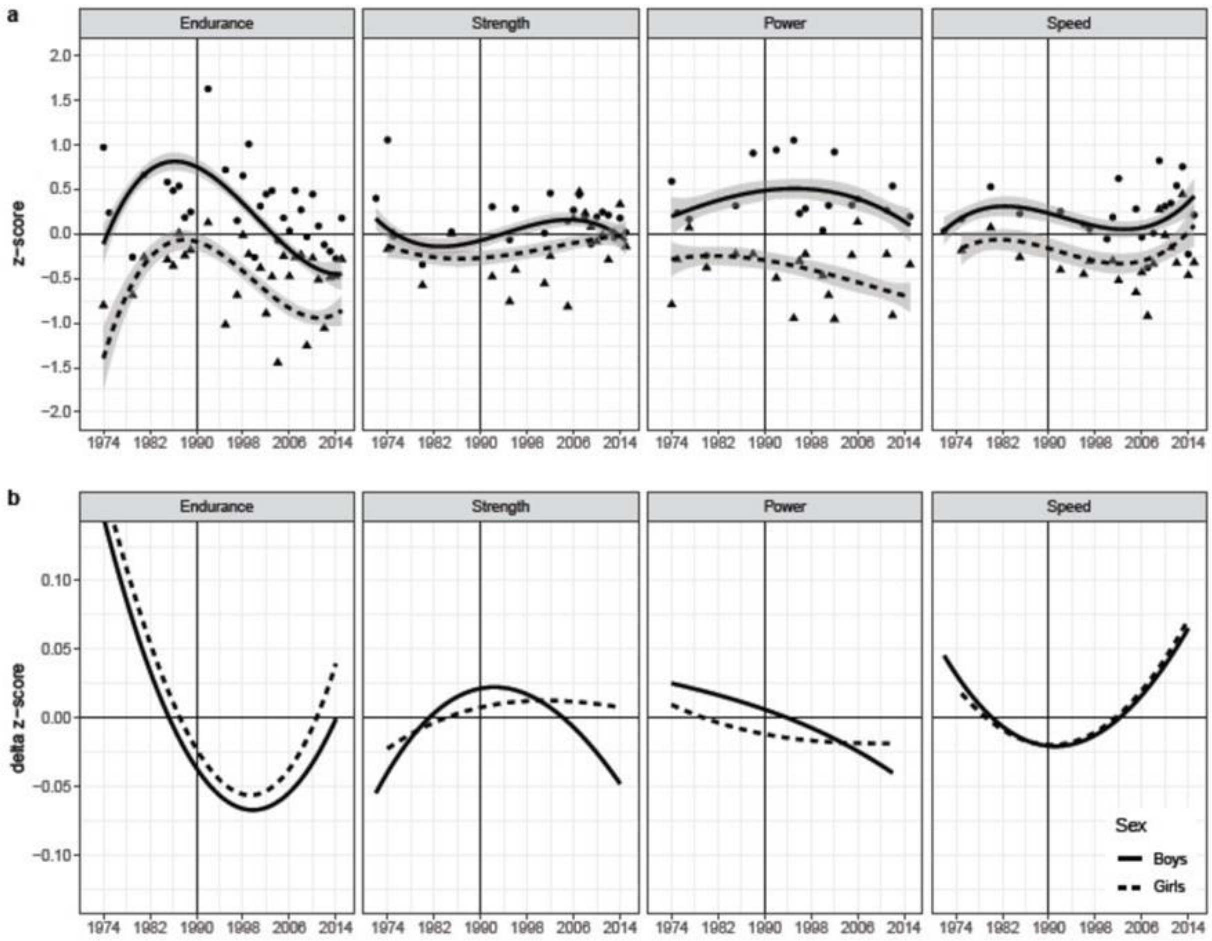

Fig. 2 a. Secular trends for cardiorespiratory endurance, relative muscle strength, proxies of muscle power, and speed for standardized scores. Positive scores indicate better physical fitness in units of the standard normal distribution (z-scores) relative to a Grand Mean of zero estimated for the year 1990. Dots represent the observed study means; the line is a third-order polynomial regression based on partial effects predicted from data and linear mixed model parameters (see Table 4); shaded areas represent 95% confidence intervals for the regression lines. b. Numeric approximations of first derivatives (delta z-score) of corresponding functions in panel a (i.e., local changes in the direction of the function). Minima and maxima in panel a correspond to zero-crossings in panel b; inflection points in cubic functions in panel a correspond to minima and maxima of first derivatives in panel b

Minima, maxima, and – for cubic Z-functions – inflection points provide critical information about secular trends. The years at which they occur are directly visible in the first derivative delta z-score of such function. They were computed with a numeric approximation depicting the instantaneous rate of change (see Figure 2b). Specifically, minima and maxima of the z score function in a-panels are at the years at which the delta z score function in b panels crosses the horizontal zero-line; inflection points of cubic z score functions are at years with minima and maxima of delta z score functions[2]. With these advanced organizers we turn to the component-specific secular trends and how they differ for boys and girls.

3.2.1 Cardiorespiratory Endurance

The LMM included 168 observations for cardiorespiratory endurance. The critical source of variance is the positive cubic trend ($b = 0.00012$, $z = 3.94$, $p < 0.001$; see Figure 2a, cardiorespiratory endurance panel). Children's cardiorespiratory endurance increased until 1986 and then decreased until around 2010–2012 with a negative inflection point in 2000. Since 2010, cardiorespiratory endurance may have stabilized or even tended to improve again (see zero-crossings and minimum in cardiorespiratory endurance panel Figure 2b for identification of years). The interaction between sex and the linear trend for years is the source for a tendency towards smaller differences between boys' and girls' cardiorespiratory fitness or a steeper decline for boys than girls ($b = -0.00796$, $z = -3.28$, $p = 0.001$).

In terms of effect magnitude (i.e., the difference between maximum and minimum z-scores across years), cardiorespiratory endurance exhibited very large secular trends with a range of about 1.5 standard deviations for both boys and girls. These effects are by far the largest of the four fitness components.

3.2.2 Relative Muscle Strength

The LMM included 156 observations for relative muscle strength. The critical source of variance is the cubic secular trend for all children (see Figure 2a, relative muscle strength panel); $b = -0.00004$, $z = -2.14$, $p = 0.032$. There are no significant effects associated with sex, but the trend is clearer for boys. An initial decline with a local minimum in 1982 is followed by an increase to a local maximum in 2006 with a positive inflection point in 1990. The magnitude of the secular trend is small; scores move consistently within a narrow 0.50 z-difference band.

3.2.3 Proxies of Muscle Power

The LMM included 164 observations for proxies of muscle power. The interaction between sex and the linear trend across years translates into shifting a very shallow negative quadratic curve ($b = -0.0006$, $z = -2.04$, $p = 0.042$) to different peaks in 1982 for girls and 1994 for boys; $b = 0.00474$, $z = 2.80$, $p = 0.005$. With 0.5, the effects are in the small-to-medium range.

[2] These relations derive from the power rule of differentiation; of relevance here are:
$f(x) = x \to f'(x) = \text{constant}$, $f(x) = x^2 \to f'(x) = 2x$, and $f(x) = x^3 \to f'(x) = 3x^2$.

3.2.4 Speed

The LMM included 164 observations for speed. Boys and girls exhibit a similar shape of their secular speed function (see also the close to identical delta z score functions in the speed panel of Figure 2b). The shape of the function is defined by a significant negative linear ($b = -0.01822$, $z = -4.90$, $p < 0.001$) and a significant positive cubic ($b = 0.00006$, $z = 4.60$, $p < 0.001$) trend. The first local peak was in 1980 (a very shallow one); the negative inflection point was in 1990; speed bottomed out in 2002 and has been rising since then. Effect sizes are about 0.50 for both groups.

3.2.5 Effect Sizes and Conditional Means

The four components of physical fitness varied qualitatively with respect to the shape of their associated secular trends and differed between boys and girls in cardiorespiratory endurance and proxies of muscle power. They also differed in effect sizes. Effect sizes were large for cardiorespiratory endurance (1.5 units of standard deviation) and small-to-medium (0.5 units of standard deviations) for the other three components. Importantly, the partial-effect functions are corrected for various sources of heterogeneity between studies as described in the following.

LMMs afford tests of heterogeneity of studies with estimates of variance components (VCs) as model parameters that account for differences between studies beyond the residual error variance. Specifically in the present data, there was no evidence for speed-related and proxies-of-muscle-power-related VCs, but there were significant VCs for cardiorespiratory endurance (0.36)[3], relative muscle strength (0.13), sex (0.18), and three linear yearly trends associated with cardiorespiratory endurance (0.03), relative muscle strength (0.013), and proxies of muscle power (0.011); the residual standard deviation was 0.25. Thus, there was considerable heterogeneity between studies and it was largest for cardiorespiratory endurance.

Data and model parameters can be used to generate predictions of conditional means at the level of studies for components of physical fitness with reliable inter-study differences. These predictions are shown in Figure 3 along with intervals of +/- 2 conditional standard deviations (~95% confidence intervals), based on the conditional variance–covariance matrices of the random effects returned by the *ranef(model, condVar=TRUE)* command of the *lme4* R package [45].

[3] Values are variances expressed as standard deviations (i.e., the square-root of the estimated VC).

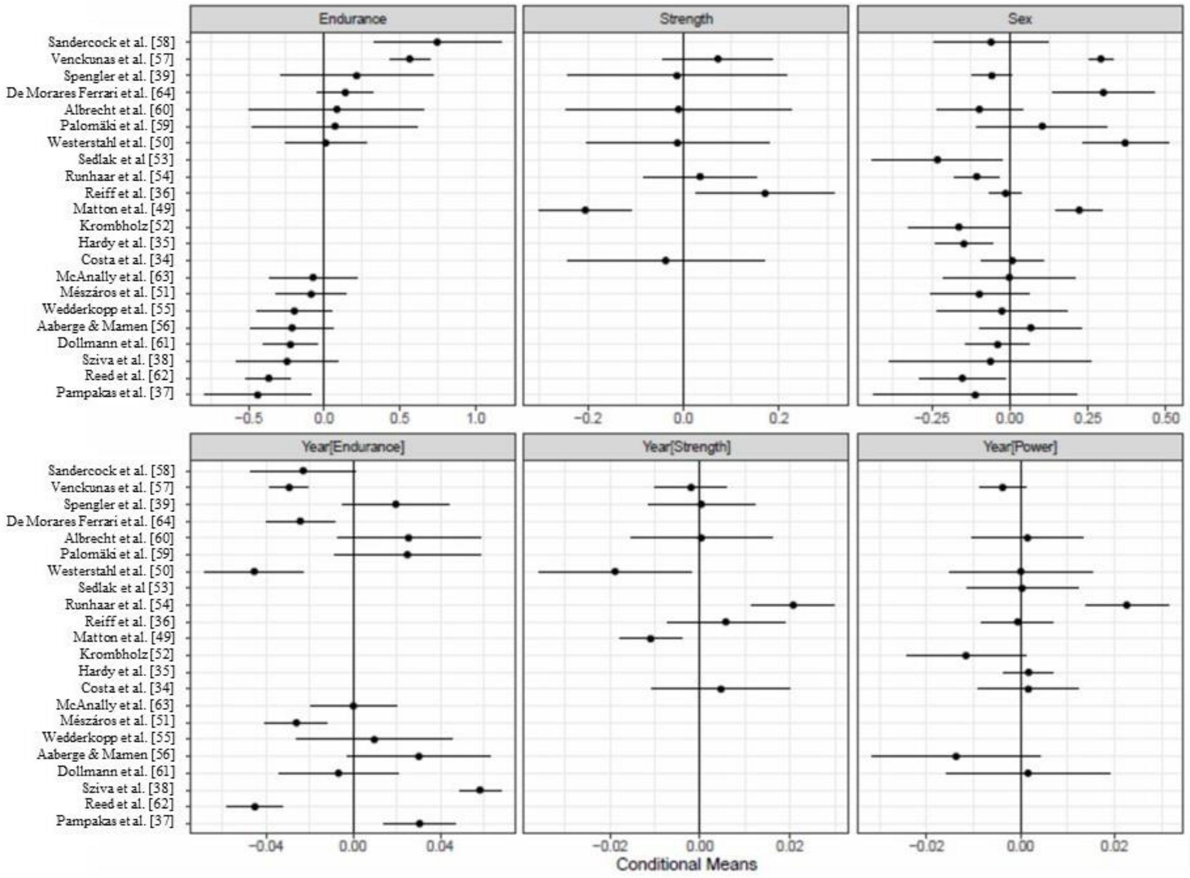

Fig. 3 Conditional means of six significant variance components in the linear mixed model

The conditional means differ from the observed means because they are shrunken towards the estimated population mean, conditional on (a) the distance of the observed from the grand mean, (b) the variance of the observed mean, and (c) the number of observations. Studies with confidence intervals completely to the left or the right of the vertical zero line (representing the grand mean), contributed significantly to heterogeneity. The major determinants of the width of the confidence intervals are the number of observations and their variance. For example, Veckunas et al. [57] contributed 48 and Sandercock et al. [58] four observations to the cardiorespiratory endurance component, accounting for the large difference between their respective intervals. Studies without an entry in a panel did not contribute observations to these components or their conditional mean was shrunk to the population mean. Taking into account this between-study heterogeneity was critical for the shape of the partial-effects functions in Figure 2a.

3.2.6 Observation-Level Model Residuals

The quality of a model fit depends on distributional characteristics of the observation-level model residuals, most importantly that they are normally distributed and homoscedastic across the range of fitted values. Figure 4 displays a few diagnostic statistics (a: LMM residuals over year by component; b: standardized LMM residuals over theoretical quantiles of the standard normal distribution; c: LMM residuals over fitted values).

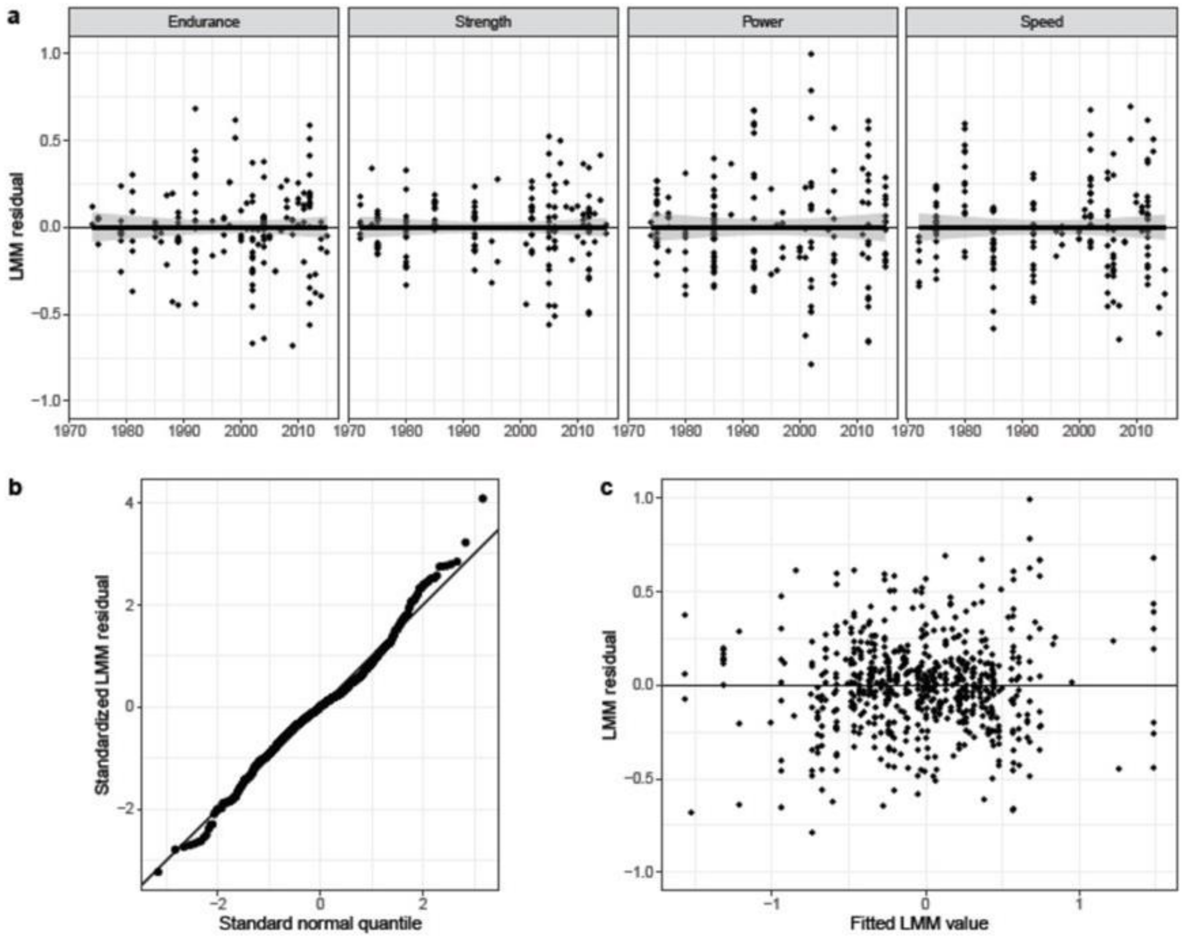

Fig. 4 Observation-level residuals of linear mixed model (LMM). a: LMM residuals over year by physical fitness component; b: standardized LMM residuals over theoretical quantiles of the standard normal distribution; c: LMM residuals over fitted values

The plots provide no evidence for violations of model assumptions. Somewhat surprisingly, there was no need for dealing with outliers. We interpret this as support for the quality of data selection, effectiveness of controlling for study heterogeneity as well as the selection of tests and their assignment to the four components of physical fitness.

4. Discussion

The aim of this systematic review was to provide an update on secular trends in selected components of physical fitness of children and adolescents aged 6–18 years. The main findings of this systematic review including 22 studies and 652 observations were (a) a large initial increase and an equally large subsequent decrease between 1986 and 2010–12 in cardiorespiratory endurance; the decrease appears to have reached a floor for all children between 2010 and 2015, but it was steeper for boys than girls, (b) a general trend towards a small increase in relative muscle strength, (c) an overall small negative quadratic trend for proxies of muscle power; and (d) a small-to-medium increase in speed since 2002.

4.1 Secular Trends and Effects of Sex for Components of Physical Fitness

In this systematic review, we provide new information on secular trends in physical fitness of children and adolescents (i.e., cardiorespiratory endurance, relative muscle strength, proxies of muscle power, and speed). There is evidence that cardiorespiratory endurance particularly declined between 1986 and 2010–12. Of note, it seems that the decline diminished during recent years for boys and girls. This finding confirms a recently published update by Tomkinson et al. [17] for cardiorespiratory endurance which showed that the international rate in cardiorespiratory endurance decline has slowed, but stabilized since the turn of the century for both boys and girls. Furthermore, Tomkinson et al. [17] reported that declines in cardiorespiratory endurance were larger for boys compared to girls which is also in line with our results.

These corresponding findings between our systematic review and the analysis of Tomkinson et al. [17] emphasize that the two reviews complement each other despite their different methodological approaches. While Tomkinson et al. [17] focused on the 20-m shuttle run, we additionally included data of time- and distance-related cardiorespiratory endurance tests. Tomkinson et al. [17] used the performance data from the 20-m shuttle run test and the equation provided by Léger and colleagues [65] to estimate $\dot{V}O_{2peak}$ from 20-m shuttle run performance for children and adolescents aged 8–19 years, irrespective of sex. In contrast, we used z-scores as a common metric for all included tests.

Tomkinson and Olds [15] outlined that secular trends in physical fitness are affected by several aspects like changed social, behavioral, physical, psychosocial, and physiological factors. Especially psychosocial factors such as motivation, one's willingness to push to maximally, and pacing strategies may have played a role, but unfortunately cannot be controlled. There are several reasons why especially fat mass and BMI may influence cardiorespiratory fitness. First, two studies reported that 40% [66]–70% [67] of the changes in cardiorespiratory endurance can be explained by changes in BMI. Second, performance in weight-bearing activities such as running over longer distances, etc. is negatively influenced by a high fat mass [28]. Cureton et al. [68] showed that gains in weight / higher fat mass significantly increased the energy cost during running. Therefore, the $\dot{V}O2max$ (expressed relative to body mass) has a lower asymptote. Third, trends in cardiorespiratory endurance (decline which was recently mitigated) coincide with international trends of overweight and obesity. The NCD Risk Factor Collaboration [69] reviewed an increasing prevalence in overweight and obesity in school-aged children worldwide between 1975 and 2016, but stated that the trend has recently flattened, albeit at a high level, especially in high-income countries.

Another factor that might affect cardiorespiratory endurance is the level of PA. Although evidence-based information about secular trends in PA are missing [70], a recently published study by Guthold et al. [4] stated that in 2016 about 80% of children and adolescents (girls: 84.7% and boys: 77.6%) worldwide did not meet the recommended 60 min at moderate-to-vigorous PA per day proposed by the WHO. However, compared with data of 2001 PA levels have been relatively stable in girls over the past 15 years (85.1%), and have slightly improved in boys (80.1%). Keeping in mind the causal interaction between PA and physical fitness [11], the low(er) levels of PA might lead to a diminished exercise stimulus and consequently to reduced cardiorespiratory endurance [66].

Our update provided information on secular trends in relative muscle strength indicating a general trend towards a small increase in fitness related to relative muscle strength. In this respect, our systematic review also complements a recently published analysis conducted by Dooley et al. [18] about secular trends in absolute muscle strength using handgrip data. They showed that the international rate of improvement in handgrip strength progressively increased between 1967 and 2017 in 2,000,000 children and adolescents aged 9–17 years living in 19 countries.

With respect to proxies of muscle power, our systematic review revealed a small negative quadratic trend and with respect to speed a small-to-medium increase in recent years. Here, we describe some differences with an earlier comprehensive review that analyzed similar measures of speed and proxies of muscle power. In fact, Tomkinson [16] summarized the existing literature on physical fitness of children and adolescents between 1958 and 2003. He included proxies of muscle power and speed data of 25,000,000 children and adolescents aged 6–19 years living in 27 countries across five geographical regions between 1958 and 2003. For proxies of muscle power, Tomkinson [16] reported a decline starting in 1985 and no evidence for differences between boys and girls. We observed a decline starting in 1995 after a previous increase for boys and a flattening for girls. According to the Tomkinson review [16], speed has flattened since about 1985 after a previous slight increase for boys and girls. In contrast, we observed an initial increase which turned to a decline at about 1982 until 2002 for boys and girls. We note that our effects are in the small-to-medium range and that we had fewer studies contributing to these two fitness components than for cardiorespiratory endurance. Tomkinson [16] considered a much larger data base including the grey literature, many more children and adolescents and many more countries. The differences may well be due to differences in methods and study design.

Compared to secular trends in cardiorespiratory endurance (1.5 SD for boys and for girls), smaller trends were observed for relative muscle strength, proxies of muscle power, and speed (0.5–1.0 SD). The reason for the observed different effects might be the varying influence of fat mass / BMI on components of physical fitness. Tomkinson et al. [28] examined to what extent fat mass / BMI influences cardiorespiratory endurance, proxies of muscle power, and speed. Fat mass / BMI correlated stronger with cardiorespiratory endurance (1600 m run) than with a proxy of muscle power (standing broad jump) or speed (50-m sprint). The correlations were r = -0.52 for cardiorespiratory endurance, r = -0.20 for proxies of muscle power, and r = -0.24 for speed [28]. Furthermore, Tomkinson et al. [28] stated that for fat-free mass, the effect is the opposite. The correlations between fat-free mass and proxies of muscle power (r = 0.59) and speed (r = 0.57) were stronger than for cardiorespiratory endurance (r = 0.21). Based on our findings and the available data, we cannot conclude whether it is a matter of fitness or fatness. When taking the work of Tomkinson et al. [28] into account, it seems that fatness plays a role particularly with cardiorespiratory endurance.

4.2 Overall Interpretation and Implications

Overall, this systematic review demonstrated that there is no generic secular trend across the different physical fitness components and years because of significant cubic secular trends in measures of cardiorespiratory endurance, relative muscle strength, proxies of muscle power, and speed. Therefore, physical fitness components must best be analyzed separately because the components showed different patterns of secular trends over time. We recommend to regularly examine secular trends according to the specific physical fitness components and the monitored time span to detect changes in physical fitness as early as possible. Schools offer a suitable setting for the assessment of physical fitness because they reach all children during their development, growth, and formation of habits.

Because of the different trends in physical fitness observed in this systematic review and the literature, we recommend to further initiatives in PA and fitness promotion for children and adolescents. More specifically, public health efforts should focus on exercising cardiorespiratory endurance and muscle strength because cardiorespiratory endurance [13] and muscle strength [8] have been reported to be positively associated with markers of health (i.e., body mass index, waist circumference, body fatness, bone mass) in children and adolescents. In addition, sufficient levels of muscle strength are a prerequisite for motor skill learning [71].

In this context, schools have a responsibility for promoting a physically active lifestyle because it is the only setting that reaches all children regardless of their socioeconomic background [72]. Therefore, structural curricula such as active recess, active education or active transportation to school, etc. are reasonable factors to reduce sedentary behavior and promote PA [73, 74]. It is crucial that parents and families are also included in these means [59]. Furthermore, the role of physical education classes need to be strengthened to enable enhanced PA and fitness levels [73–75]. Several studies verify the positive impact of resistance training [76–80] and endurance training [81, 82] on children's physical fitness. Based on this research, muscle strengthening and aerobic exercises can be included in the curricula as several studies [83–85] indicated that well-structured and well-designed physical education lessons can have a positive impact on physical fitness of children and adolescents. In addition, Faigenbaum and Bruno [86] stated that children and adolescents who are regularly engaged in PA (e.g., through physical education, etc.) were more eager to develop complex movements and sport skills (e.g., in a sport club or free play).

4.3 Limitations

This systematic review has some limitations. First, biological maturity or anthropometric data (e.g., body mass or body height) were not considered in the analysis because this information were not available in most of the included studies. There is strong evidence that physical fitness is related to biological maturity as more mature children and adolescents outperform less mature children and adolescents in physical fitness [87]. Furthermore, studies [88, 89] indicated a trend towards earlier maturation (age of puberty onset) over the last few centuries which is accompanied by an increase in physical fitness which may have influenced and altered secular trends. Future studies on secular trends in physical fitness of children and adolescents should always report the biological maturity status if possible. Second, every included study used different study designs (tests, instruments, methods, environmental conditions, etc.) for measuring physical fitness which resulted in inter-study variance especially for cardiorespiratory endurance and relative muscle strength. To better evaluate secular trends in physical fitness, future studies are advised to use identical test procedures among representative samples of children and adolescents worldwide [90]. Third, secular trends were mainly reported for high-income countries such as the United States, Australia, and Europe. Future studies should also be conducted in middle- and low-income countries using simple, cheap, valid, and reliable physical fitness measures. Fourth, the included studies are usually not representative of the population of their countries. Future studies should collect country-specific representative data. Fifth, all included studies reported mean values of secular trends. However, trends in mean values might be systematically biased if there have been concurrent trends in skewness whereas median values will not be biased. Future studies should also relate to measures of centrality (e.g., medians), variability (e.g., standard deviations), and asymmetry (e.g., skewness). Sixth, we do not claim that our systematic review was exhaustive; grey literature, project reports or unpublished work were not considered. Finally, an assessment of risk of bias was not conducted. Accordingly, future studies should include risk of bias assessment in their analysis.

5. Conclusions

This systematic review documents a large initial increase and an equally large subsequent decrease for cardiorespiratory endurance starting in 1986 and lasting until 2010–12. The decrease appears to have reached a floor for all children between 2010 and 2015. Measures of relative muscle strength showed a general trend towards a small increase. Measures of proxies of muscle power indicated an overall small negative quadratic trend. For measures of speed, a small-to-medium increase was observed since 2002. Because of the different trends in physical fitness, we recommend to further initiatives in PA and fitness promotion for children and adolescents. More specifically, public health efforts should focus on exercising cardiorespiratory endurance to prevent adverse health effects (i.e., overweight, obesity) and muscle strength to lay a foundation for motor skill learning. Further, studies are needed that examine dose–response relations for muscle and bone strengthening exercises in children and adolescents.

Compliance with Ethical Standards

Funding

Open Access funding enabled and organized by Projekt DEAL. No sources of funding were used to assist in the preparationon this article.

Conflicts of interest/competing interests

Thea Fühner, Reinhold Kliegl, Fabian Arntz, Susi Kriemler, and Urs Granacher declare that they have no conflicts of interest relevant to the content of this systematic review.

Availability of data and material

The R script is available in Electronic Supplementary Material Appendix S2. Data from the included studies are available in the Electronic Supplementary Material Appendix S3. Furthermore, the R script and the data set from all included studies are available through the Open Science Framework (OSF): https://osf.io/rs7x2/?view_only=090b2e63c23e4e719f766f6791ed1a89

Code availability

Not applicable

Author contributions

TF, RK, SK, and UG: made substantial contributions to conception and design; TF and FA: contributed to data collection; RK and TF carried out data analysis; TF, RK, SK, and UG: interpreted the data; TF: wrote the first draft of the manuscript and all authors were involved in revising it critically for important intellectual content; all authors provide final approval of the version to be published and agreed to be accountable for all aspects of the work.

References

1. World Health Organization. Global recommendations on physical activity for health. Geneva: World Health Organization; 2010.
2. Kalman M, Inchley J, Sigmundova D, Iannotti RJ, Tynjälä JA, Hamrik Z, et al. Secular trends in moderate-to-vigorous physical activity in 32 countries from 2002 to 2010: a cross-national perspective. Eur J Public Health. 2015;25 Suppl 2:37–40. doi:10.1093/eurpub/ckv024.
3. Hallal PC, Andersen LB, Bull FC, Guthold R, Haskell W, Ekelund U. Global physical activity levels: surveillance progress, pitfalls, and prospects. The Lancet. 2012;380:247–57. doi:10.1016/S0140-6736(12)60646-1.
4. Guthold R, Stevens GA, Riley LM, Bull FC. Global trends in insufficient physical activity among adolescents: a pooled analysis of 298 population-based surveys with 1·6 million participants. Lancet Child Adolesc Health 2019. doi:10.1016/S2352-4642(19)30323-2.
5. Faigenbaum AD, Myer GD. Exercise deficit disorder in youth: play now or pay later. Curr Sports Med Rep. 2012;11:196–200.
6. Ortega FB, Ruiz JR, Castillo MJ, Sjöström M. Physical fitness in childhood and adolescence: a powerful marker of health. Int J Obes. 2008;32:1–11.
7. Poitras VJ, Gray CE, Borghese MM, Carson V, Chaput JP, Janssen I, et al. Systematic review of the relationships between objectively measured physical activity and health indicators in school-aged children and youth. Appl Physiol Nutr Metab. 2016;41:197-239. doi:10.1139/apnm-2015-0663.
8. García-Hermoso A, Ramírez-Campillo R, Izquierdo M. Is muscular fitness associated with future health benefits in children and adolescents? A systematic review and meta-analysis of longitudinal studies. Sports Med. 2019;49:1079–94. doi:10.1007/s40279-019-01098-6.
9. Telama R, Yang X, Leskinen E, Kankaanpää A, Hirvensalo M, Tammelin T, et al. Tracking of physical activity from early childhood through youth into adulthood. Med Sci Sports Exerc. 2014;46:955–62. doi:10.1249/MSS.0000000000000181.

10. Wrotniak BH, Epstein LH, Dorn JM, Jones KE, Kondilis VA. The relationship between motor proficiency and physical activity in children. Pediatrics. 2006;118:e1758-65. doi:10.1542/peds.2006-0742.

11. Larsen LR, Kristensen PL, Junge T, Rexen CT, Wedderkopp N. Motor Performance as predictor of physical activity in children: The CHAMPS Study-DK. Med Sci Sports Exerc. 2015;47:1849–56. doi:10.1249/MSS.0000000000000604.

12. Caspersen CJ, Powell KE, Christenson GM. Physical activity, exercise, and physical fitness: definitions and distinctions for health-related research. Public Health Rep. 1985;100:126–31.

13. Mintjens S, Menting MD, Daams JG, van Poppel MNM, Roseboom TJ, Gemke RJBJ. Cardiorespiratory fitness in childhood and adolescence affects future cardiovascular risk factors: a systematic review of longitudinal studies. Sports Med. 2018;48:2577–605. doi:10.1007/s40279-018-0974-5.

14. Pate RR. Health-related measures of children's physical fitness. J School Health. 1991;61:231–3.

15. Tomkinson GR, Olds TS. Secular changes in pediatric aerobic fitness test performance: the global picture. In: Tomkinson GR, Olds TS, editors. Pediatric fitness: secular trends and geographic variability. Basel: Karger; 2007. p. 46–66.

16. Tomkinson GR. Global changes in anaerobic fitness test performance of children and adolescents (1958-2003). Scand J Med Sci Sports. 2007;17:497–507. doi:10.1111/j.1600-0838.2006.00569.x.

17. Tomkinson GR, Lang JJ, Tremblay MS. Temporal trends in the cardiorespiratory fitness of children and adolescents representing 19 high-income and upper middle-income countries between 1981 and 2014. Br J Sports Med. 2019;53:478–86. doi:10.1136/bjsports-2017-097982.

18. Dooley FL, Kaster T, Fitzgerald JS, Walch TJ, Annandale M, Ferrar K, et al. A systematic analysis of temporal trends in the handgrip strength of 2,216,320 children and adolescents between 1967 and 2017. Sports Med. 2020;50:1129–44. doi:10.1007/s40279-020-01265-0.

19. World Health Organization. Global action plan on physical activity 2018-2030: more active people for a healthier world. Geneva: World Health Organization; 2018.

20. Moher D, Liberati A, Tetzlaff J, Altman DG. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. PLoS Med. 2009;6:1–6. doi:10.1371/journal.pmed.1000097.g001.

21. Liberati A, Altman DG, Tetzlaff J, Mulrow C, Gøtzsche PC, Ioannidis JPA, et al. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration. PLoS Med. 2009;6:1-28. doi:10.1371/journal.pmed.1000100.

22. Eisenmann JC, Malina RM. Secular trend in peak oxygen consumption among United States youth in the 20th century. Am J Hum Biol. 2002;14:699–706. doi:10.1002/ajhb.10084.

23. Macfarlane D, Tomkinson GR. Evolution and variability in fitness test performance of Asian children and adolescents. In: Tomkinson GR, Olds TS, editors. Pediatric fitness: secular trends and geographic variability. Basel: Karger; 2007. p. 143–167.

24. Malina RM, Katzmarzyk PT. Physical activity and fitness in an international growth standard for preadolescent and adolescent children. Food Nutr Bull. 2006;27:295–313.

25. Malina RM. Physical fitness of children and adolescents in the United States: status and secular change. In: Tomkinson GR, Olds TS, editors. Pediatric fitness: secular trends and geographic variability. Basel: Karger; 2007. p. 67–90.

26. Shephard RJ. Fitness of Canadian children: range from traditional inuit to sedentary city dwellers, and assessment of secular changes. In: Tomkinson GR, Olds TS, editors. Pediatric fitness: secular trends and geographic variability. Basel: Karger; 2007. p. 91–103.

27. Tomkinson GR, Léger LA, Olds TS, Cazorla G. 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. Sports Med. 2003;33:285–300.

28. Tomkinson GR, Hamlin MJ, Olds TS. Secular changes in anaerobic test performance in Australasian children and adolescents. Pediatr Exerc Sci. 2006;18:314–28.

29. Tomkinson GR, Olds TS. Secular changes in aerobic fitness test performance of Australasian children and adolescents. In: Tomkinson GR, Olds TS, editors. Pediatric fitness: secular trends and geographic variability. Basel: Karger; 2007. p. 168–182.

30. Tomkinson GR, Olds TS, Kang SJ, Kim DY. Secular trends in the aerobic fitness test performance and body mass index of Korean children and adolescents (1968-2000). Int J Sports Med. 2007;28:314–20. doi:10.1055/s-2006-924357.

31. Tomkinson GR, Macfarlane D, Noi S, Kim DY, Wang Z, Hong R. Temporal changes in long-distance running performance of Asian children between 1964 and 2009. Sports Med. 2012;42:267–79.

32. Tomkinson GR, Olds TS, Borms J. Who are the eurofittest? In: Tomkinson GR, Olds TS, editors. Pediatric fitness: secular trends and geographic variability. Basel: Karger; 2007. p. 104–128.

33. Jürimäe T, Volbekiene V, Jürimäe J, Tomkinson GR. Changes in eurofit test performance of Estonian and Lithuanian children and adolescents (1992-2002). In: Tomkinson GR, Olds TS, editors. Pediatric fitness: secular trends and geographic variability. Basel: Karger; 2007. p. 129–142.

34. Costa AM, Costa MJ, Reis AA, Ferreira S, Martins J, Pereira A. Secular trends in anthropometrics and physical fitness of young Portuguese school-aged children. Acta Med Port. 2017;30:108–14. doi:10.20344/amp.7712.

35. Hardy LL, Merom D, Thomas M, Peralta L. 30-year changes in Australian children's standing broad jump: 1985-2015. J Sci Med Sport. 2018;21:1057–61. doi:10.1016/j.jsams.2018.04.005.

36. Reiff GG, Dixon WR, Jacoby D, Ye GX, Spain CG, Hunsicker PA. The president's council on physical fitness and sports 1985: national school population fitness survey. Washington: President's Council on Physical Fitness and Sports; 1986.

37. Pampakas P, Meszaros Z, Kiraly T, Szmodis MB, Szakaly Z, Zsidegh M, Meszaros J. Longitudinal differences and trends in body fat and running endurance in Hungarian primary schoolboys. Anthropol Anz. 2008;66:317–26.

38. Sziva A, Mészáros Z, Kiss K, Mavroudes M, Ng N, Mészáros J. Longitudinal differences in running endurance and body mass index - a 25-year comparison. Acta Physiol Hung. 2009;96:359–68. doi:10.1556/APhysiol.96.2009.3.9.

39. Spengler S, Rabel M, Kuritz AM, Mess F. Trends in motor performance of first graders: a comparison of cohorts from 2006 to 2015. Front Pediatr 2017. doi:10.3389/fped.2017.00206.

40. Wickham H, Averick M, Bryan J, Chang W, McGowan L, François R, et al. Welcome to the tidyverse. J Open Source Softw. 2019;4:1686. doi:10.21105/joss.01686.

41. Hohenstein S, Kliegl R. remef: remove partial effects. 2020. https://github.com/hohenstein/remef/.

42. Lüdecke D. sjPlot: data visualization for statistics in social science. 2020. https://cran.r-project.org/web/packages/sjPlot/.

43. Wilke CO. Cowplot: streamlined plot theme and plot annotations for 'ggplot2'. 2019. https://cran.r-project.org/web/packages/cowplot/index.html.

44. Bolker B, Robinson D. broom.mixed: tidying methods for mixed models. 2020. https://cran.r-project.org/web/packages/broom.mixed/index.html.

45. Bates D, Mächler M, Bolker B, Walker S. Fitting linear mixed-effects models using lme4. J Stat Softw. 2015b;67:1–48. doi:10.18637/jss.v067.i01.

46. R Core Team. R: a language and environment for statistical computing. 2020. https://www.R-project.org/.

47. Bates D, Kliegl R, Vasishth S, Baayen H. Parsimonious mixed models. 2015. http://arxiv.org/abs/1506.04967.

48. Matuschek H, Kliegl R, Vasishth S, Baayen H, Bates D. Balancing type I error and power in linear mixed models. J Mem Lang. 2017;94:305–15. doi:10.1016/j.jml.2017.01.001.

49. Matton L, Duvigneaud N, Wijndaele K, Philippaerts R, Duquet W, Beunen G, et al. Secular trends in anthropometric characteristics, physical fitness, physical activity, and biological maturation in Flemish adolescents between 1969 and 2005. Am J Hum Biol. 2007;19:345–57. doi:10.1002/ajhb.20592.

50. Westerstahl M, Barnekow-Bergkvist M, Hedberg G, Jansson E. Secular trends in body dimensions and physical fitness among adolescents in Sweden from 1974 to 1995. Scand J Med Sci Sports. 2003;13:128–37.

51. Mészáros J, Mahmoud O, Szabo T. Secular trend and motor performance scores in Hungarian school-boys. Facta Univ Ser Phys Educ Sport. 1999;1:43–9.

52. Krombholz H. Haben sich motorische Leistungen von 3- bis 7-jährigen Jungen und Mädchen im Zeitraum von 1973 bis 2001 verschlechtert? Z Sportpsychol. 2011;18:161–71. doi:10.1026/1612-5010/a000060.

53. Sedlak P, Pařízková J, Daniš R, Dvořáková H, Vignerová J. Secular Changes of adiposity and motor development in Czech preschool children: lifestyle changes in fifty-five year retrospective study. Biomed Res Int. 2015:1–9. doi:10.1155/2015/823841.

54. Runhaar J, Collard DCM, Singh AS, Kemper HCG, van Mechelen W, Chinapaw M. Motor fitness in Dutch youth: differences over a 26-year period (1980-2006). J Sci Med Sport. 2010;13:323–8. doi:10.1016/j.jsams.2009.04.006.

55. Wedderkopp N, Froberg K, Hansen HS, Andersen LB. Secular trends in physical fitness and obesity in Danish 9-year-old girls and boys: odense school child study and Danish substudy of the european youth heart study. Scand J Med Sci Sports. 2004;14:150–5.

56. Aaberge K, Mamen A. A comparative study of fitness levels among norwegian youth in 1988 and 2001. Sports. 2019;7:1–11. doi:10.3390/sports7020050.

57. Venckunas T, Emeljanovas A, Mieziene B, Volbekiene V. Secular trends in physical fitness and body size in Lithuanian children and adolescents between 1992 and 2012. J Epidemiol Commun H. 2017;71:181–7. doi:10.1136/jech-2016-207307.

58. Sandercock GRH, Ogunleye A, Voss C. Six-year changes in body mass index and cardiorespiratory fitness of English schoolchildren from an affluent area. Int J Obes. 2015;39:1504–7. doi:10.1038/ijo.2015.105.

59. Palomäki S, Heikinaro-Johansson P, Huotari P. Cardiorespiratory performance and physical activity in normal weight and overweight Finnish adolescents from 2003 to 2010. J Sports Sci. 2015;33:588–96. doi:10.1080/02640414.2014.951874.

60. Albrecht C, Hanssen-Doose A, Bös K, Schlenker L, Schmidt S, Wagner M, et al. Motorische Leistungsfähigkeit von Kindern und Jugendlichen in Deutschland: Eine 6-Jahres-Kohortenstudie im Rahmen der Motorik-Moduls (MoMo). Sportwissenschaft. 2016;46:294–304. doi:10.1007/s12662-016-0421-4.

61. Dollmann J, Olds T, Norton K, Stuart D. The Evolution of fitness and fatness in 10-11-year-old Australian schoolchildren: changes in distributional characteristics between 1985 and 1997. Pediatr Exerc Sci. 1999;11:108–21.

62. Reed KE, Warburton DER, Whitney CL, McKay HA. Secular changes in shuttle-run performance: a 23-year retrospective comparison of 9- to 11-year-old children. Pediatr Exerc Sci. 2006;18:364–73.

63. McAnally HM, Reeves LM, Sligo JL, Hancox RJ. Intergenerational changes in adolescents´ physical fitness and weight in New Zealand. N Z Med J. 2018;131:16–28.

64. De Moraes Ferrari, G. L., Bracco MM, Rodrigues-Matsudo VK, Fisberg M. Cardiorespiratory fitness and nutritional status of schoolchildren: 30-year evolution. J Pediatr. 2013;89:366–73. doi:10.1016/j.jped.2012.12.006.

65. Léger LA, Mercier D, Gadoury C, Lambert J. The multistage 20 metre shuttle run test for aerobic fitness. J Sports Sci. 1988;6:93–101. doi:10.1080/02640418808729800.

66. Olds T, Dollmann J. Are changes in distance-run performance of Australian children between 1985 and 1997 explained by changes in fatness? Pediatr Exerc Sci. 2004;16:201–9.

67. Albon HM, Hamlin MJ, Ross JJ. Secular trends and distributional changes in health and fitness performance variables of 10-14-year-old children in New Zealand between 1991 and 2003. Br J Sports Med. 2010;44:263–9. doi:10.1136/bjsm.2008.047142.

68. Cureton KJ, Sparling PB, Evans BW, Johnson SM, Kong UD, Purvis JW. Effect of experimental alterations in excess weight on aerobic capacity and distance running performance. Med Sci Sport Exer. 1978;10:194–9.

69. NCD Risk Factor Collaboration. Worldwide trends in body-mass index, under-weight, overweight, and obesity from 1975 to 2016: a pooled analysis of 2416 population-based measurement studies in 128,9 million children, adolescents, and adults. Lancet. 2017;390:2627–42.

70. Ekelund U, Tomkinson G, Armstrong N. What proportion of youth are physically active? Measurement issues, levels and recent time trends. Br J Sports Med. 2011;45:859–65. doi:10.1136/bjsports-2011-090190.

71. Granacher U, Lesinski M, Busch D, Muehlbauer T, Prieske O, Puta C. Effects of resistance training in youth athletes on muscular fitness and athletic performance. a conceptual model for long-term athlete development. Front Physiol. 2016;164:1–14.

72. United Nations Educational, Scientific and Cultural Organization. International charter of physical education, physical activity and sport; 2015. Paris: United Nations Educational, Scientific and Cultural Organization; 2015.

73. Meyer U, Schindler C, Zahner L, Ernst D, Hebestreit H, van Mechelen W, et al. Long-term effect of a school-based physical activity program (KISS) on fitness and adiposity

in children: a cluster-randomized controlled trial. PLoS ONE. 2014;9:e87929. doi:10.1371/journal.pone.0087929.

74. Kriemler S, Meyer U, Martin E, van Sluijs EMF, Andersen LB, Martin BW. Effect of school-based interventions on physical activity and fitness in children and adolescents: a review of reviews and systematic update. Br J Sports Med. 2011;45:923–30. doi:10.1136/bjsports-2011-090186.

75. Meyer U, Roth R, Zahner L, Gerber M, Puder JJ, Hebestreit H, Kriemler S. Contribution of physical education to overall physical activity. Scand J Med Sci Sports. 2013;23:600–6. doi:10.1111/j.1600-0838.2011.01425.x.

76. Drenowatz C, Greier K. Resistance training in youth - benefits and characteristics. J Biomed. 2018;3:32–9. doi:10.7150/jbm.25035.

77. Faigenbaum AD, Myer GD. Pediatric resistance training: benefits, concerns, and program design considerations. Curr Sports Med Rep. 2010;9:161–8.

78. Myers AM, Beam NW, Fakhoury JD. Resistance training for children and adolescents. Transl Pediatr. 2017;6:137–43. doi:10.21037/tp.2017.04.01.

79. Faigenbaum AD, MacDonald JP. Dynapenia: it's not just for grown-ups anymore. Acta Paediatr. 2017;106:696–7. doi:10.1111/apa.13797.

80. Lesinski M, Herz M, Schmelcher A, Granacher U. Effects of resistance training on physical fitness in healthy children and adolescents: an umbrella review. Sports Med 2020. doi:10.1007/s40279-020-01327-3.

81. Armstrong N, Tomkinson G, Ekelund U. Aerobic fitness and its relationship to sport, exercise training and habitual physical activity during youth. Br J Sports Med. 2011;45:849–58. doi:10.1136/bjsports-2011-090200.

82. Armstrong N. Aerobic fitness and training in children and adolescents. Pediatr Exerc Sci. 2016;28:7–10. doi:10.1123/pes.2016-0004.

83. Granacher U, Muehlbauer T, Doerflinger B, Strohmeier R, Gollhofer A. Promoting strength and balance in adolescents during physical education: effects of a short-term resistance training. J Strength Cond Res. 2011;25:940–9.

84. Faigenbaum AD, Bush JA, McLoone RP, Kreckel MC, Farrell A, Ratamess NA, Kang J. Benefits of strength and skill-based training during primary school physical education. J Strength Cond Res. 2015;29:1255–62.

85. Fu Y, Burns RD, Constantino N, Fitzsimmons J, Zhang P. Effect of the resistance exercise on elementary school students' physical fitness. SSEJ. 2019;17:468–75. doi:10.1007/s42978-019-0022-7.

86. Faigenbaum AD, Bruno LE. A fundamental approach for treating pediatric dynapenia in kids. ACSMs Health Fit J. 2017;21:18-24.

87. Jones MA, Hitchen PJ, Stratton G. The importance of considering biological maturity when assessing physical fitness measures in girls and boys aged 10 to 16 years. Ann Hum Biol. 2000;27:57–65. doi:10.1080/030144600282389.

88. Bellis MA, Downing J, Ashton JR. Adults at 12? Trends in puberty and their public health consequences. J Epidemiol Commun H. 2006;60:910–1. doi:10.1136/jech.2006.049379.

89. Herman-Giddens ME. Recent data on pubertal milestones in United States children: the secular trend toward earlier development. Int J Androl. 2006;29:241-6; discussion 286-90. doi:10.1111/j.1365-2605.2005.00575.x.

90. Fraser BJ, Blizzard L, Tomkinson GR, Lycett K, Wake M, Burgner D, et al. The great leap backward: changes in the jumping performance of Australian children aged 11-12-years between 1985 and 2015. J. Sports Sci. 2019;37:748–54. doi:10.1080/02640414.2018.1523672.

Publication II

AGE AND SEX EFFECTS IN PHYSICAL FITNESS COMPONENTS OF 108,295 THRID GRADERS INCLUDING 515 PRIMARY SCHOOLS AND 9 COHORTS

Thea Fühner[1], Urs Granacher[1], Kathleen Golle[1], Reinhold Kliegl[1]

[1] University of Potsdam, Research Focus Cognition Sciences, Division of Training and Movement Science, Potsdam, Germany

Reference

Fühner, T., Granacher, U., Golle, K., & Kliegl, R. (2021b). Age and sex effects in physical fitness components of 108,295 third graders including 515 primary schools and 9 cohorts. *Scientific Reports, 11*, 17566. https://doi.org/10.1038/s41598-021-97000-4

The final publication is available at:

https://www.nature.com/articles/s41598-021-97000-4

Age and sex effects in physical fitness components of 108,295 third graders including 515 primary schools and 9 cohorts

Thea Fühner[1], Urs Granacher[1], Kathleen Golle[1], Reinhold Kliegl[1*]

[1]Division of Training and Movement Sciences, Research Focus Cognition Sciences, University of Potsdam, Potsdam, Germany

[*]Corresponding author and contact details

Prof. Reinhold Kliegl

Division of Training and Movement Sciences

Research Focus Cognitive Sciences

University of Potsdam

Am Neuen Palais 10, Building 12

14469 Potsdam

Germany

Tel: +49 331 977 1303 or +49 160 94 76 22 12

E-mail address: reinhold.kliegl@uni-potsdam.de

ORCID ID: orcid.org/0000-0002-0180-8488

Abstract

Children's physical fitness development and related moderating effects of age and sex are well documented, especially boys' and girls' divergence during puberty. The situation might be different during prepuberty. As girls mature approximately two years earlier than boys, we tested a possible convergence of performance with five tests representing four components of physical fitness in a large sample of 108,295 eight-year old third-graders. Within this single prepubertal year of life and irrespective of the test, performance increased linearly with chronological age, and boys outperformed girls to a larger extent in tests requiring muscle mass for successful performance. Tests differed in the magnitude of age effects (gains), but there was no evidence for an interaction between age and sex. Moreover, "physical fitness" of schools correlated at r = 0.48 with their age effect which might imply that "fit schools" promote larger gains; expected secular trends from 2011 to 2019 were replicated.

Introduction

Children's development of physical fitness as well as the effects of moderating variables such as age and sex are well documented[1–7], especially boys' and girls' divergence during puberty starting late in the twelfth and tenth year of life, respectively[8]. The situation might be different during prepuberty. Girls mature visibly about two years earlier than boys, but sex hormones rise already much starting at eight years of age[9–11]. If the early rise of sex hormones relates to body composition, the question arises whether there is evidence for faster development of girls than boys in this year, leading to a convergence of performance in physical fitness during the transition from pre-puberty to puberty?

A longitudinal study on the development of different fitness components highlights sex-specific performance trajectories in youth aged 9 to 12 years[6]. This study was a precursor project of the present cross-sectional study and tested four components of physical fitness in four annual assessments including 240 children[6]. The main panels in Figure 1 illustrate sex- and age-differential development for cardiorespiratory endurance, coordination, speed, and power (assessed separately for lower [powerLOW] and upper [powerUP] limbs). The five tests used for the assessment are described in the figure caption.

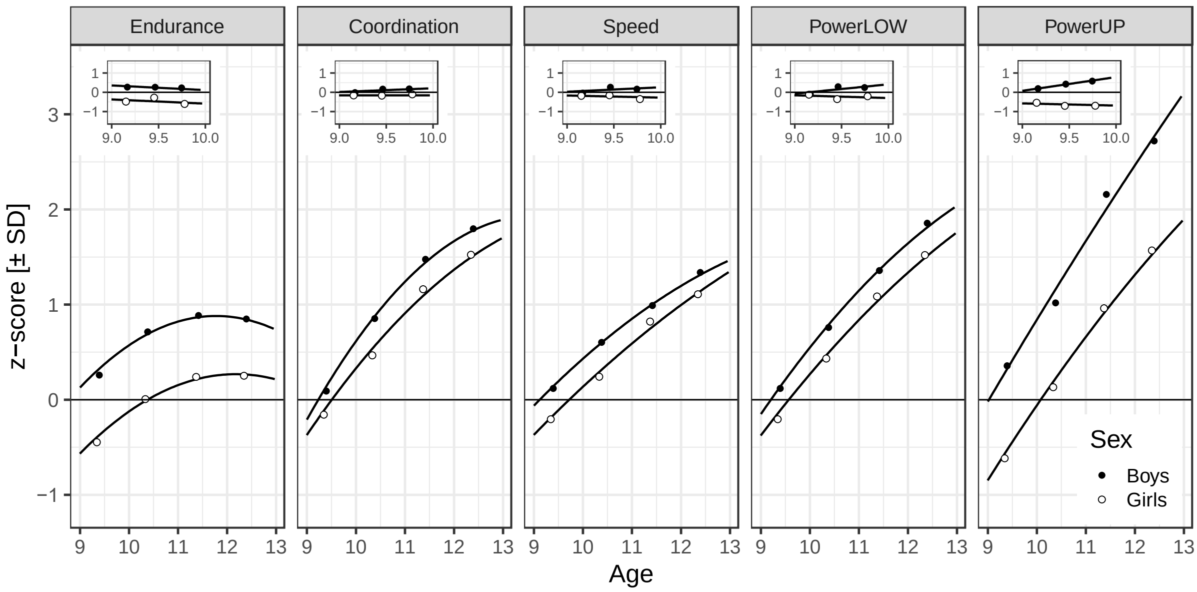

Figure 1. Physical fitness curves of a longitudinal sample of 240 German boys (closed circles) and girls (open circles) followed from age 9 to 12 years for endurance = cardiorespiratory endurance (i.e., 9 min run test), coordination (i.e., running in a star like pattern), speed (i.e., 50-m linear sprint test), powerLOW = power of lower limbs (i.e., triple hop test), and powerUP = power of upper limbs (i.e., ball push test). The insets show for each test score the regression on age for the first assessment when children were between 9.00 and 9.99 years old. Also shown are the means for groups of boys and girls binned into three age groups (i.e., 9.00-9.33; 9.34-9.66; 9.67-9.99). Error bands are 95% CIs. Data are from Golle et al.[6].

Speed, powerLOW, and powerUP follow a mostly linear trajectory, whereas cardiorespiratory endurance and coordination are characterized by a curvilinear development[6]. These age-related differences can mostly be attributed to growth (increasing body mass and body height) and maturation during childhood and adolescence[1]. For instance, while increased body mass may have a positive impact on ball push test performance, it may negatively influence performance in tests such as the 6 or 9 min run tests which afford the continuous acceleration of the body[12]. Furthermore, the fact that boys significantly outperform girls in these tests[6] is due to differences in muscle mass favouring boys; there is evidence that prepuberal boys have on average 3.7% larger muscle mass compared with girls[1,13]. In line with this argument, Golle et al.[6] observed larger sex differences in fitness tests demanding muscle mass (e.g., powerUP) compared with tests of motor coordination (e.g., running in a starlike pattern).

The longitudinal study of Golle et al.[6] and other cross-sectional studies[4,5,7,14,15] showed performance increases for all components of physical fitness over childhood and adolescents. However, it may come as a surprise that in the longitudinal study of Golle et al.[6] none of the cross-sectional age differences and none of the interactions with sex were significant within the respective assessment years - neither when aggregated over tests nor when tested for individual tests; the insets in Figure 1 show these non-significant trends and the means for three age groups for the first assessment of tests when children were between 9.00 and 9.99 years (see

Supplement B for details of a reanalysis). However, the absence of evidence for cross-sectional age differences and their interactions with sex within a single year of life at a prepubertal development stage must not be taken as evidence for their absence[16], because statistical power may not have been sufficiently large.

Our study included a very large representative sample of 108,295 eight-year old third graders. This large sample allowed us to zoom into a cross-sectional short-term ontogenetic window. In this single year of life, we expected to detect age differences that were not significant in the insets of Figure 1. Indeed, we expected that the component differences in developmental gains (i.e., the differences in slopes) will be in general agreement with those in the main panel of Figure 1 because gains should be larger (favouring older children) in tests that do not demand continuous acceleration of the body as with the ball push test vs the 6 min run test[12]. Similarly, as in Figure 1 (both in the main panel and in the insets) the test-related sex effects (i.e., the differences between the lines) will be larger, the more muscle mass is required to perform a physical fitness test[1,13] (favouring boys) and less for tests involving motor coordination[17]. In other words, in general, we expected to anticipate the long-term ontogenetic longitudinal trends across four years with short-term ontogenetic cross-sectional trends within a single year for younger children in the third grade.

There is one exception to this expectation of agreement between short-term and long-term ontogenetic trends: The only significant interaction between age and sex was reported for powerUP (see right panel in Figure 1) indicating divergence between boys and girls. If physical fitness components carry an early prepubertal signal in eight-year olds, then we should observe convergence of scores because girls will benefit earlier than boys from the rise of sex hormones[9,11].

Cross-sectional analysis has been criticized for good reasons and in general preference is given to longitudinal analysis, mostly because only the latter delivers information about intraindividual growth[6]. However, we propose that, when the focus is on short-term ontogenesis, that is on changes within, not between years of life, a cross-sectional design is probably the only option to determine development-related gains because this design circumvents practice effects, a necessary consequence of repeated testing of physical fitness within a year. Another problem of cross-sectional designs are cohort effects (i.e., age-correlated cultural change). However, cohort effects are certainly negligible for children who attend the same grade but differ in age only within the same year of life. Thus, with the exception of loss of information about interindividual differences in intraindividual change, a cross-sectional design is more suitable to determine how within-year developmental profiles differ between health-related (e.g., cardiorespiratory endurance)[18] and skill-related components of physical fitness (e.g., power, speed, coordination)[18].

Just as there are individual differences in physical fitness between children, there are also differences between the over 500 schools in how much they implement programs that facilitate gains in the development of physical fitness[19]. Explanatory hypotheses about these differences are beyond the scope of this article, but our use of a linear mixed model for statistical inference affords exploratory tests for their presence and adjusts test statistics for school-related sources of variance. Finally, as data collection for this cross-sectional study occurred annually from 2011 to 2019, secular trends are another source of variance in scores that must be taken into account. Here we expected that the results will be in line with recent original research and meta-analyses and show a decline in cardiorespiratory endurance and an increase in speed[20–22].

In summary, the hypotheses about age- and sex-related differences in physical fitness tests were tested with a linear mixed model (LMM) that afforded the simultaneous consideration of children, schools, and cohorts as random factors and the estimation of variance components and correlation parameters for (a) test scores, (b) effects of contrasts between the tests, and (c) in the case of schools and cohorts also effects of age-related gains and sex differences. These model parameters serve primarily as measures of statistical control for the fixed effect estimates but may also yield substantive insights about the dynamics of development. For example, for correlations of scores, we expected the usual positive manifold between the tests, but for correlations of effects (i.e., the contrasts between tests) no directed hypotheses were formulated given the usually low reliability associated with difference scores. Although age and sex are between-child factors, they are also both within-school and within-cohort factors, providing us with the opportunity to detect reliable variance components and correlation parameters for these factors. No directed hypotheses were formulated for schools. However, secular trends are well documented, and we expected to replicate them.

Results

Overview

Table 1 displays statistics for fixed effects of age (linear) and sex as well as their interactions with the four test contrasts for LMM m2. Test-specific z-transformations eliminated main effects of contrasts H1 to H4 (all $z \leq 0.99$). Neither the age x sex interaction nor any of the interactions of this term with the four test contrasts were significant (all $|z| \leq 1.74$). Adding quadratic trends of age and their associated interactions to the model, did not significantly contribute to goodness of fit; $\chi^2(10) = 9.17$, $p = 0.52$. None of the age x sex interaction was significant when tested separately for the five tests (all $|z| \leq 1.04$; see Supplement A for details about both control LMMs).

Age-related gains

Figure 2 displays both the gains in physical fitness with age and boys' higher scores than girls' in each of the five physical fitness tests. The parallel lines in each panel also visualize the lack of significant evidence for age x sex as well as age x sex x test interactions.

Table 1 Fixed-effect estimates of linear mixed model

Source of variance	Fixed-effect estimates	Standard error	z-values	Pr (> |z|)
Main effects				
Grand mean (intercept)	-0.038	0.011	-3.57*	< 0.001
H1: coordination vs. endurance	0.019	0.024	0.76	0.445
H2: speed vs. coordination	-0.033	0.033	-1.01	0.313
H3: powerLOW vs. speed	0.033	0.033	0.99	0.323
H4: powerUP vs. powerLOW	0.005	0.020	0.24	0.808
Age (linear)	0.271	0.009	31.60*	< 0.001
Sex	0.413	0.005	86.52*	< 0.001
Age (linear) x Sex	0.002	0.014	0.14	0.888
Age (linear) x Component				
H1: coordination vs. endurance	0.216	0.012	18.44*	< 0.001
H2: speed vs. coordination	-0.069	0.011	-6.22*	< 0.001
H3: powerLOW vs. speed	-0.007	0.010	-0.69	0.491
H4: powerUP vs. powerLOW	0.307	0.012	25.87*	< 0.001
Sex x Component				
H1: coordination vs. endurance	-0.260	0.007	-39.03*	< 0.001
H2: speed vs. coordination	0.078	0.006	12.35*	< 0.001
H3: powerLOW vs. speed	0.066	0.006	11.24*	< 0.001
H4: powerUP vs. powerLOW	0.296	0.007	43.84*	< 0.001
Age (linear) x Sex x Component				
H1: coordination vs. endurance	0.0393	0.023	1.74	0.082
H2: speed vs. coordination	-0.014	0.021	-0.63	0.527
H3: powerLOW vs. speed	-0.020	0.020	-0.99	0.324
H4: powerUP vs. powerLOW	0.025	0.023	1.07	0.282

H1 to H4 = hypothesis 1 to 4, endurance = cardiorespiratory endurance (i.e., 6 min run test), coordination = star run test, speed = 20-m linear sprint test, powerLOW = power of lower limbs (i.e., standing long jump test), powerUP = power of upper limbs (i.e., ball push test), * = z-value > 3.0, linear mixed model random factors: cohorts (9), schools (515), children (108,295), observations = 525,126 (missing = 3%). For estimates of variance components and correlation parameters see Table 3.

Despite the small age range, the differences were large and visible for all five physical fitness tests; the overall linear trend for age was significant ($z = 31.60$). The LMM tested the interactions of age with the four test contrasts, that is whether slopes in neighbouring panels (averaged across sex) were parallel. The left two panels (H1) and the right two panels (H4) show that this is clearly not the case – even without a statistical test. Indeed, three of four expected interactions were significant (see second block of Table 1): the age slope was larger for coordination than cardiorespiratory endurance (H1; z = 18.44), larger for coordination than speed (H2; z = -6.22), and larger for powerUP than powerLOW (H4; z = 25.87). The difference between age slopes for powerLOW and speed (H3) was not significant (z = -0.69).

Counter to this profile of differences, Figure 2 which is based on observed scores suggests that speed gain is larger than powerLOW gain. Indeed, this contrast was significant for LMM m1, that is as long as VCs for cohort-related differences between tests were not in the LMM. The large differences in secular trends are shown in Figure 4 (below); adjusting for them in the LMM yielded the reported partial effects.

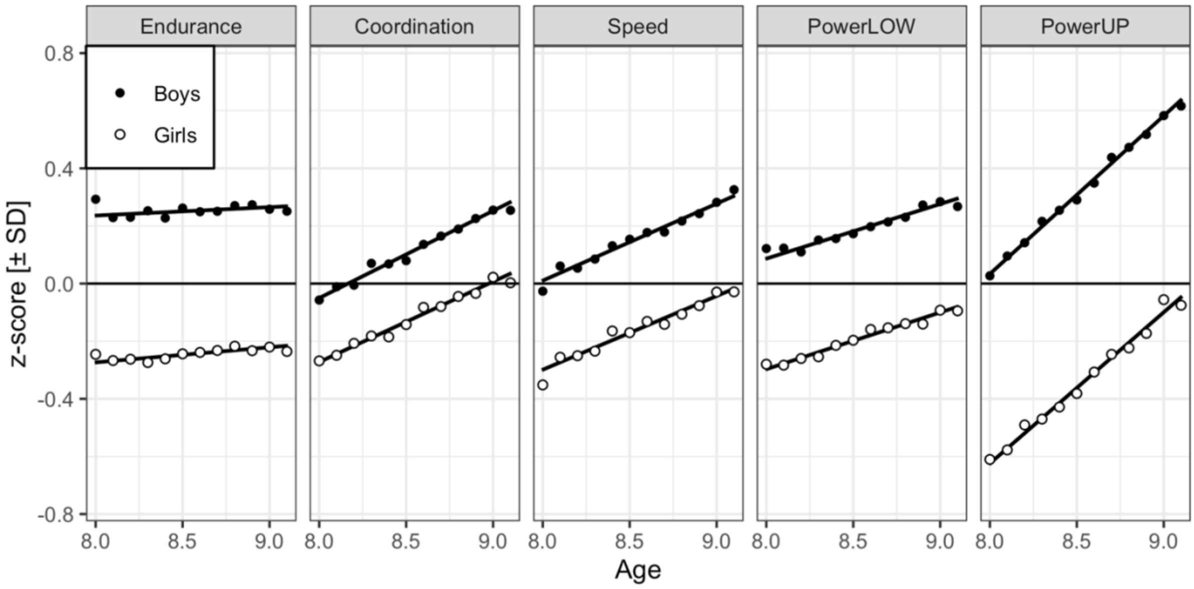

Figure 2. Performance differences between 8.0 und 9.0 years by sex in the five physical fitness tests presented as z-transformed data computed separately for each test. Endurance = cardiorespiratory endurance (i.e., 6 min run test), Coordination = star run test, Speed = 20-m linear sprint test, PowerLOW = power of lower limbs (i.e., standing long jump test), PowerUP = power of upper limbs (i.e., ball push test), SD = standard deviation. Points are binned observed child means; lines are simple regression fits to the observations; 95% confidence intervals for means ≈ 0.05 are not visible.

Sex-related effects

The difference between lines in Figure 2 displays the expected differences between boys and girls for the performance in the five physical fitness tests; the overall sex effect was significant with z = 86.52. The third block of Table 1 lists statistics for the interactions between sex and the tests contrasts. All interaction terms were significant and in agreement with a priori expectations. Boys performed better than girls on cardiorespiratory endurance than coordination (H1; z = -39.03), better on speed than coordination (H2; z = 12.35), better on powerLOW than speed (H3; z = 11.24), and better on powerUP than powerLOW (H4; z = 43.84). The magnitude of these interactions is shown in Figure 2 in the differences between the parallel lines for boys and girls for neighboring panels (averaged over age).

Variance components and correlation parameters

Test scores

Table 2 lists estimates of VCs and CPs for the five test scores from a re-parameterized version of LMM `m2` with the same goodness of fit and the same estimates for fixed-effects. The test-related VCs were large for children (0.72 to 0.78), of medium-size for schools (0.24 to 0.37), and small for cohorts (0.03 to 0.06). VCs for the age-related gains (slopes 0.10) and the sex effect (0.05) were also small for schools. It is noteworthy that the differences between schools in the age-related gain of their children is larger than the differences between cohorts.

Table 2 Variance components, correlation parameters, and zero-order correlations for test scores

	VC	CP \ r						
Component		**End**	**Coord**	**Speed**	**PowerLOW**	**PowerUP**	**Age**	**Sex**
Child								
Endurance	0.74	1.00	**0.37**	**0.41**	**0.42**	0.23		
Coordination	0.76	**0.56**	1.00	**0.44**	**0.44**	0.32		
Speed	0.76	**0.61**	**0.67**	1.00	**0.52**	0.32		
PowerLOW	0.78	**0.59**	**0.65**	**0.77**	1.00	0.37		
PowerUP	0.72	0.25	0.45	0.43	0.50	1.00		
School								
Endurance	0.32	1.00	**0.33**	**0.33**	0.38	0.17	0.15	0.05
Coordination	0.37	**0.35**	1.00	**0.32**	0.37	0.25	0.14	0.03
Speed	0.31	**0.34**	**0.33**	1.00	0.46	0.29	0.10	0.06
PowerLOW	0.25	**0.37**	**0.39**	**0.43**	1.00	0.30	0.12	0.06
PowerUP	0.24	0.19	0.18	0.28	0.26	1.00	0.13	-0.04
Age	0.10	0.44	0.34	0.28	0.32	0.21	1.00	0.10
Sex	0.05	0.12	-0.05	0.11	0.19	-0.04	0.25	1.00
Cohort								
Endurance	0.05							
Coordination	0.04							
Speed	0.06							
PowerLOW	0.04							
PowerUP	0.03							

End = cardiorespiratory endurance (i.e., 6 min run test), Coord = star run test, Speed = 20-m linear sprint test, PowerLOW = power of lower limbs (i.e., standing long jump test), PowerUP = power of upper limbs (i.e., ball push test), VC = variance component, CP \ r = correlation parameter \ zero-order correlation; linear mixed model correlation parameters are shown in black and corresponding pairwise zero-order correlations in light grey for children (top) and schools (middle). Theoretically relevant correlations discussed in the text are set in **bold**. VC for Residual = 0.53. VCs and CPs are based on full set of data; ZOC correlations are based on subsets of data (Child: 96,529 children from 512 schools; School: 93,661 children from 421 schools).

In the first block of Table 2, CPs between tests scores for children are shown in black and the corresponding zero-order correlations (ZOCs) are shown in light grey. The ZOCs are based on 96,529 children with complete test scores; they came from 512 different schools. In the second block, we list corresponding results for schools. For this analysis, we also added the criterion that a school had to report complete data from more than 30 boys and 30 girls to ensure stable estimation of age and sex effects within schools. This criterion left us with 421 schools, 93,661 children, and 468,305 scores.

There were three noteworthy patterns of results. First, as expected, all child- and school-related CPs and ZOCs between test scores were positive. Thus, the five tests represent a latent construct "physical fitness" both for differences between children and for differences between schools. This was also supported by a random-effects principal component analyses (rePCA) of the two orthogonal random-effect structures. The first principal component (PC1) loadings ranged from 0.49 to 0.35 for the child-related PC1 and from 0.45 to 0.29 for the school-related PC1, accounting for 65% and 38% of the respective variances (see Supplement A for details).

Second, the tests did not correlate equally highly with each other. Most notably, the child-related CPs of cardiorespiratory endurance, coordination, speed, and powerLOW correlated very highly between 0.56 and 0.77, but their correlations with powerUP were distinctly smaller (CPs: 0.25 to 0.50). This observation holds also for the other three correlation matrices, but overall correlations were smaller (see Table 2 for all CPs).

Again, this interpretation was supported by rePCAs (see Supplement A for details). The smallest loading on child- and school-related PC1s was obtained for powerUP (0.35, 0.29). Moreover, for children the loadings for the second PC2s (15%) represented the difference between cardiorespiratory endurance (0.53) and powerUP (-0.84). Similarly, for schools the third PC3 (12%) represented the difference between the average of cardiorespiratory endurance (0.31) and coordination (0.47) and powerUP (-0.71). We take this PC2/PC3-based difference score as support for the hypothesis that powerUP favoured heavier children for who strength of arms may mask reduced cardiorespiratory endurance and coordination (see Discussion).

Third, child-related CPs (Table 2, black numbers) were larger than child-based ZOCs (Table 2, light grey numbers). This was a rather striking pattern because one might expect the opposite given that ZOCs were confounded with large effects of age and sex. Conversely, CPs were larger despite adjustment for sex and age differences in the fixed effects and for differences due to schools and cohorts in the random-effect structure of the LMM. The reason for the result is LMM-based shrinkage of conditional means of the units of the random factors in the direction of the GM. Thus, the entire data set was used to "correct" unreliable observations, also called "borrowing strength" to improve predictions. Due to this shrinkage correction for unreliability, CPs revealed the latent relations between measures much more clearly than ZOCs.

Effects of test contrasts

In the random-effect structure of LMM m2, estimates were returned for child-, school-, and cohort related VCs for GM and the four test contrasts; VCs of age and sex were also estimated for school. CPs for child and school reflect correlations between the contrasts (i.e., effect correlations). The results are shown in Table 3. As in Table 2, CPs are shown in black and corresponding ZOCs are light grey.

VCs for test contrasts were larger (0.47 to 0.72) for children and somewhat smaller, but still highly reliable (0.30 to 0.40) for schools, especially when compared to VCs estimated for school-related age (0.10) and sex (0.05) effects, and especially when compared to cohort-related effects (0.04 to 0.09). CPs and ZOCs of effects are smaller than CPs and ZOCs based on test scores because, with the exception of those involving GM, they are all based on difference scores.

There were two results of theoretical relevance. First, there was a negative CP for rGM.H4 for children (-0.31): If we invert the difference score to convert to a positive correlation, then large values of GM correspond to a large difference between powerLOW and powerUP. In other words, the larger powerLOW relative to powerUP, the larger is the expected value for GM. Thus, in line with the special status of powerUP reported above, powerUP is better thought of as an adjustment or sharpening of physical fitness as indicated by PowerLOW than as a genuine indicator of physical fitness by itself. The corresponding ZOC was only -0.13.

Table 3 Variance components, correlation parameters, and zero-order correlations for test-related contrasts

	VC	CP \ r						
Effects		**GM**	**H1**	**H2**	**H3**	**H4**	**Age**	**Sex**
Child								
Grand Mean	0.60	1.00	0.03	0.03	0.02	**-0.13**		
H1: coordination vs. endurance	0.67	0.11	1.00	-0.51	-0.01	0.05		
H2: speed vs. coordination	0.58	0.06	-0.51	1.00	-0.46	-0.07		
H3: powerLOW vs. speed	0.47	0.04	0.01	-0.36	1.00	-0.40		
H4: powerUP vs. powerLOW	0.72	**-0.31**	0.15	-0.20	-0.23	1.00		
School								
Grand Mean	0.20	1.00	0.08	-0.06	-0.11	-0.19	**0.19**	0.05
H1: coordination vs. endurance	0.40	0.14	1.00	-0.55	0.00	0.05	0.00	-0.01
H2: speed vs. coordination	0.40	-0.14	-0.59	1.00	-0.50	-0.02	-0.05	0.02
H3: powerLOW vs. speed	0.30	-0.13	0.02	-0.50	1.00	-0.38	0.00	-0.00
H4: powerUP vs. powerLOW	0.30	-0.19	-0.04	0.06	-0.38	1.00	0.00	-0.09
Age (linear)	0.10	**0.48**	-0.03	-0.11	-0.01	-0.10	1.00	0.10
Sex	0.05	0.09	-0.14	0.13	0.05	-0.19	0.25	1.00
Cohort								
Grand Mean	0.02							
H1: coordination vs. endurance	0.05							
H2: speed vs. coordination	0.08							
H3: powerLOW vs. speed	0.09							
H4: powerUP vs. powerLOW	0.04							

H1 to H4 = hypothesis 1 to 4, endurance = cardiorespiratory endurance (i.e., 6 min run test), coordination = star run test, speed = 20-m linear sprint test, powerLOW= power of lower limbs (i.e., standing long jump test), powerUP = power of upper limbs (i.e., ball push test), VC = sqrt (variance component), CP \ r = correlation parameter \ zero-order correlation; linear mixed model correlation parameters are shown in black and corresponding pairwise zero-order correlations in light grey for children (top) and schools (middle). Theoretically relevant correlations discussed in the text are set in **bold**. VC for Residual = 0.54. VCs and CPs are based on full set of data; ZOC correlations are based on subsets of data (Child: 96,529 children from 512 schools; School: 93,661 children from 421 schools).

Second, we note three large negative CPs (r H1.H2, r H2.H3, and r H3.H4). However, these correlations are ambiguous because the contrasts had a test in common (i.e., coordination is part of H1 and H2; speed is part of H2 and H3; powerLOW is part of H3 and H4).

Third, the largest effect CP was observed between age and GM for schools (+0.48) suggesting that "fitter" schools promote more developmental change across the school year. Figure 3A displays a visualization of the CP using a scatterplot of the conditional means of age-related gain over conditional means of GMs of physical fitness for the 515 schools.

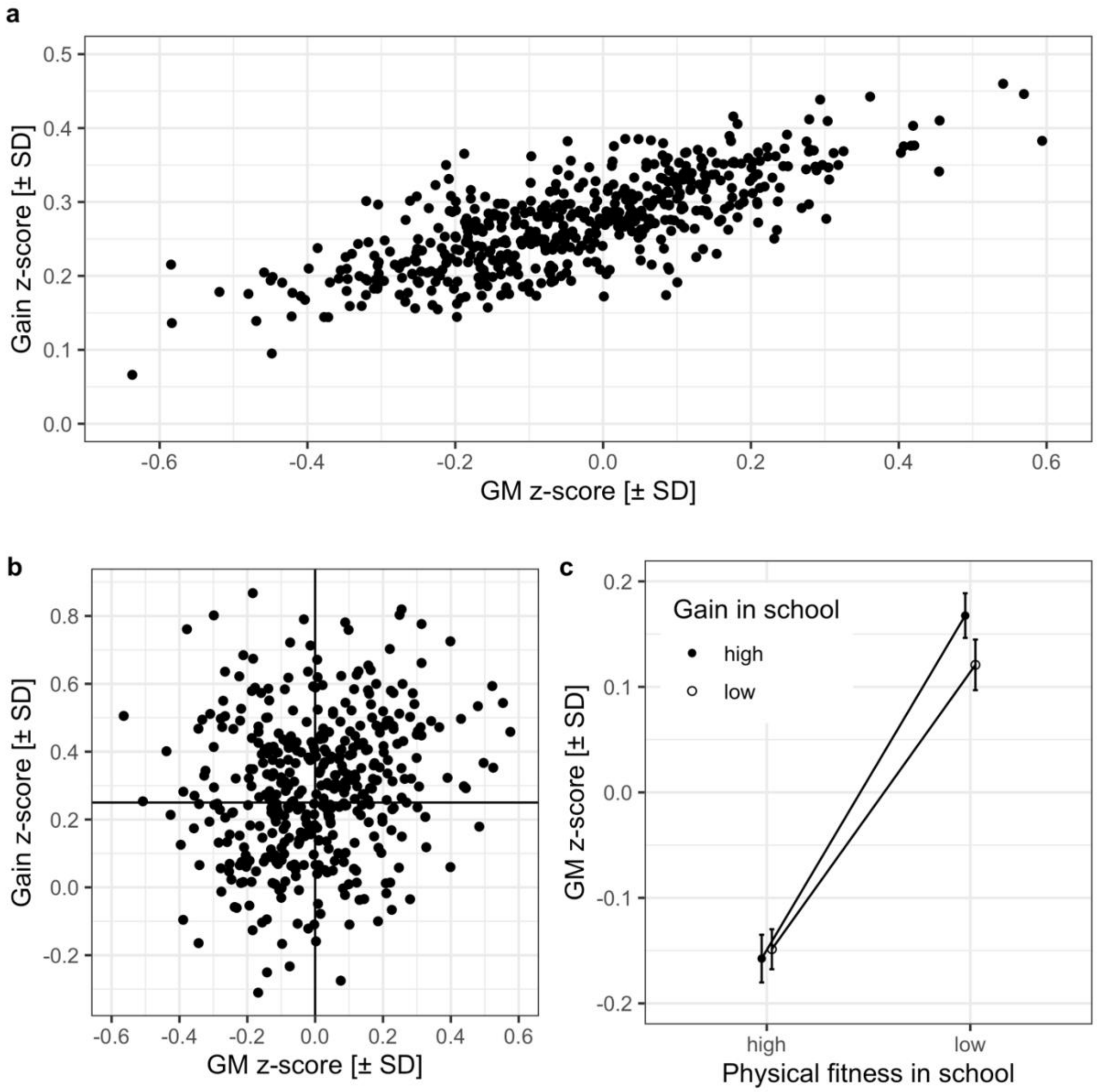

Figure 3. **(a)** Visualization of correlation parameter between Grand Mean (GM) and yearly gains (age effect) of 515 schools using conditional means resulting from shrinkage correction of observed data with LMM parameters. **(b)** Scatterplot of observed GMs and yearly gains (within-school age slopes) for 421 schools reporting data from more than 30 boys and 30 girls. **(c)** After sorting schools into groups of high vs. low physical fitness (split at 0) and high vs. small gain-rate (split at the gain of +0.25), the interaction corresponding to the CP becomes visible; error bars are 95% CIs. SD = standard deviation.

The scatterplot is not identical with the CP, in fact the correlation is 0.75, because conditional means are "predictions" of the school age effect and GM using the school data and all model parameters to correct for unreliability in the scores. Indeed, as shown in Figure 3B, there is no evidence for this relation in the uncorrected scatterplot corresponding to the ZOC of 0.19. A significant CP corresponds to a simple interaction in the data and this interaction can be visualized by sorting schools into those of high and low physical fitness (split at z-score = 0) and those with a high and small gain for their children in the third grade (split at the age slope of +0.25). In an ANOVA of schools' observed mean physical fitness values the interaction between these two post-hoc grouping factors was significant; $F(1, 417) = 4.22$, $MSe = 0.012$, $p < 0.05$. Of course, from a correlation we cannot infer the direction of causality. "Fitter schools" (e.g., schools offering extracurricular sport-related activities) may facilitate gains in children's fitness. Alternatively, if we assume that fitter children gain more in a year, then a school's high fitness as well as the associated large gain could be the result of being attended by fitter children (e.g., due to the school's location in a high-SES region).

Cohort-related variance components

The small, but significant VCs related to the random factor cohort indicate that there were reliable test x cohort interactions; they are shown in Figure 4. Across the nine years from 2011 to 2019 there was a performance decline for cardiorespiratory endurance and an increase for speed. The other three components exhibit an initial increase followed by a decline of performance in recent years.

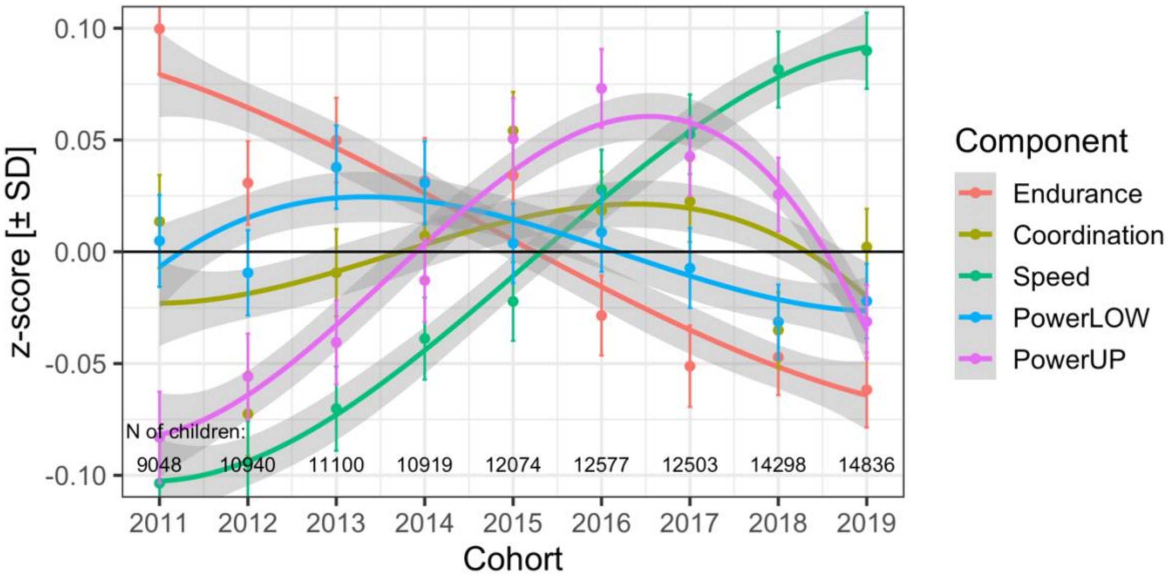

Figure 4. Cohort-related change of components of physical fitness. Points are observed means with 95% CIs. Lines are third-order polynomial trends fitted to children's scores along with 95% error. Note the much smaller range of the y-axis (i.e., from -0.10 to +0.10) compared to age effects shown in Figure 2. Endurance = cardiorespiratory endurance (i.e., 6 min run test), Coordination = star run test, Speed = 20-m linear sprint test, PowerLOW = power of lower limbs (i.e., standing long jump test), PowerUP = power of upper limbs (i.e., ball push test).

Goodness-of-fit statistics and model residuals

Additional details about the LMM analyses are documented in Supplement A which also contains information about the control LMMs. Despite their complexity, all models converged without problems and there was no evidence of estimates of parameters at their boundaries. Overparameterization was observed only for the most complex LMM m4. Thus, with this exception, the LMMs were supported by the data. Finally, we carried out residual-based diagnostics (e.g., q-q plot, standardized residuals over fitted values, etc.) for the reference LMM m2 with CPs for effects (Table 1, Table 3). These tests did not reveal any problems.

Discussion

The aim of this study was to examine short-term ontogenetic cross-sectional developmental differences in physical fitness for five tests tapping health- and skill-related components of physical fitness in a large sample of 108,295 German eight-year old children. Even in a single prepubertal year of life (1) performance increases linearly with chronological age in all physical fitness tests, (2) boys outperform girls in all physical fitness tests with sex differences being larger for tests requiring muscle mass and being smaller for tests requiring motor coordination, (3) the tests differ, mostly as expected, in the size of age and sex effects, (4) four of the five tests represent a common construct (i.e., correlate strongly positively with each other) – the exception is the ball push test that requires powerUP, (5) "physically fit schools" apparently promote more developmental gains within a year (but this is only correlational evidence) and (6) diverging secular trends for cardiorespiratory endurance (negative) and speed (positive) are in agreement with other research and meta-analyses. Furthermore, there was no evidence for an interaction of age and sex – with each other or with the test contrasts despite an abundance of statistical power. Counter to folklore, a very large data set (i.e., 108,295 subjects and 525,126 observations) did not “automatically” render everything significant! Given this statistical power, we are strongly inclined to interpret the absence of evidence for interactions as evidence of absence of interactions for these five tests of physical fitness[16].

Boys outperformed girls in all four physical fitness components (cardiorespiratory endurance, coordination, speed, power [LOW/UP]). This is in line with other studies reporting normative values[2–5,14,23–26]. For instance, Tambalis et al.[5] reported that boys aged 6 to 18 years showed significantly better performances for cardiorespiratory endurance (i.e., 20 m shuttle run test), powerLOW (i.e., standing long jump test), and agility (i.e., 10 x 5 m agility shuttle run test) compared to girls. Furthermore, De Miguel-Etayo et al.[3] described that boys aged 6 to 10 years significantly outperformed girls in cardiorespiratory endurance (i.e., 20 m shuttle run test), speed (i.e., 40 m sprint test), and powerLOW (i.e., standing long jump test).

What are the reasons for the observed large test-specific sex differences? Sex-related differences in body composition appear to be likely candidates to account for the observed findings. The larger sex effect in powerUP compared to powerLOW can be explained by a better proportion of strength relative to body mass especially in the upper limbs in boys compared to girls[27,28]. The large difference in cardiorespiratory endurance can be explained by physiological factors such as boys’ better mechanical efficiency and fractional utilisation of oxygen[2,29]. Furthermore, muscle mass[1] and muscle cross-sectional area[13] favour boys especially in physical fitness tests that recruit muscle mass. Beside these anthropometric factors and physiological demands, sociocultural aspects may also explain the sex difference. Haywood and Getchell[30] reported that girls usually participate in sports that require balance and flexibility (e.g., gymnastics, figure skating) compared to boys who rather participate in strength-related activities.

The sex effect was also significantly stronger for powerLOW than for speed. PowerLOW is determined much more by muscle mass where boys usually outperform girls[2,3,5]. In contrast, speed is less influenced by muscle mass than by motor coordination where sex differences are comparatively small[31] or were not found at all[17]. Therefore, the sex effect in powerLOW might be larger than in speed. Obviously, the demand of coordination relative to power and cardiorespiratory endurance is even larger in the star run test than in the 20-m linear sprint test[32–34] and this could be a reason why the sex effect is smaller for the star run test than for speed.

The observed decrease of sex differences in performances with increased demands for coordination seems to be plausible. Coordination draws on many different central (e.g., brain) and peripheral sites (e.g., motor units) within the nervous system[35]. The more a test engages the brain, the less relevant sex is a performance limiting factor. In summary, the decrease of size of sex effects across tests can most likely be explained by a decrease in the relevance of muscle mass (favouring boys over girls)[1] and an increase in the relevance in motor coordination (associated with small or no sex effects)[17,31].

Within their ninth year of life, older children significantly outperform younger ones in all five fitness tests. This was expected[3,5,14], but there were no significant gains for a sample of 240 children in the tenth year of life (see insets of Figure 1). The absence of evidence for nonlinear relations between performance and chronological age as well as the similarity of gains for boys and girls are quite remarkable (see Figure 2), especially in light of what long-term ontogenetic longitudinal research reveals when children are a few years older. Obviously, the current test battery does not detect the onset of puberty in the ninth year of life! Nevertheless, the different age-related grains for tests are compatible with what is known about the development of basic physiological parameters: Growth of body mass and height positively influence the performance in skill-related physical fitness components of coordination, speed and power. In contrast, increases in body mass have a negative impact on performance in tests such as cardiorespiratory endurance (e.g.., 6 min run test) where the own body mass has to be accelerated continuously. The smaller age effect for cardiorespiratory endurance compared to skill-related components to the remaining physical fitness components is in line with a study conducted by de Miguel-Etayo et al.[3]. The authors could not find significant age differences in cardiorespiratory endurance (i.e., 20 m shuttle run test) but amongst other in powerLOW (i.e., standing long jump test) and strength (i.e., absolute handgrip strength test) in a sample of children aged 6 to 10 years. Similar, Viru et al.[36] stated that an accelerated improvement in cardiorespiratory endurance occurs at the ages of 11 to 15 years in boys and 11 to 13 years in girls.

As far as the differential age effects between the components of physical fitness are concerned, tests of coordination, speed, and powerLOW share the highest correlations among the five tests. These three tests share the relevance of muscle mass yielding power, but they differ in the relevance of coordination. As mentioned above, the sex effect within these three tests (i.e., star run test < 20-m linear sprint test < standing long jump test; see differences between lines in Figure 2) is in line with their ranking on coordination. For age-related gains within the school year, the partial effects yielded star run test > 20-m linear sprint test and standing long jump test, corresponding (roughly) to their ranking on motor coordination. The special status of powerUP (i.e., ball push test) is not only evident with respect to its lower correlations with other tests, but also with respect to the size of the age effect – by far the largest of the five tests (see Figure 2).

Obviously, the performance in powerUP was influenced by factors other than physical fitness. We propose that body mass contributes to performance in tests that assess powerUP. Heavier children usually have a higher muscle mass compared to normal weighted children[37]. While a higher muscle mass positively influences performances in non-weight-bearing tests (e.g., ball push test), it negatively influences performances in weight-bearing tests (e.g., standing long jump test) because the body mass has to be accelerated in contrast to non-weight-bearing tests[12]. Therefore, tests for the assessment of powerUP favour heavier children for whom strength of arms and trunk may mask reduced cardiorespiratory endurance and coordination. Thus, a high score for powerUP may be more indicative of a lack of overall physical fitness because it does not measure physical fitness to the same degree than the other four tests (i.e., small correlations with the other four tests and GM). Our results suggest that the best indicator of physical fitness is the average of the first four tests (cardiorespiratory endurance, coordination, speed, and powerLOW). Analyses including body mass (not measured in the present study) could support the interpretation of the special status powerUP in the assessment of physical fitness.

The linear mixed model included school as a second random factor, supported the estimation of correlation parameters for test scores and age effects, and revealed a strong correlation (r = 0.48) between the age effect and overall physical fitness. The direction of causality is not clear, but in line with Hattie[19] and García-Hermoso et al.[38]. The results are in agreement with the hypothesis that schools differ in how much they qualitatively promote the development of children's physical fitness[38]. The development of physical fitness depends on the quality of the physical education lessons and especially on the quality of the physical education teacher teaching physical education[38]. These school differences in the age effect of physical fitness within the third grade were strong enough to yield reliable differences in the overall physical fitness of their children.

Our study is not without limitations. First, the five tests do not cover all components of physical fitness[18] and there are alternative tests for each component. Although our results do not necessarily generalize to other components like muscle strength, muscle endurance, or balance[18], they represent those for whom an early detection of puberty was most likely. Second, from a cross-sectional study we cannot know whether the linear gains for tests hold at the individual level or are the result of averaging over individual differences in non-linear growth curves. Obviously, high-density monitoring within a year would be desirable, but such longitudinal data are not without their own problems. For example, learning effects could be reduced due to the completion of at least three familiarization sessions[39] to separate learning effects from growth. However, motivational factors may also play a role. The longitudinal cardiorespiratory endurance data in Figure 1 suggest no further growth or even a decline in performance for 12 year old children[6]. This is obviously not in agreement with what we know about the objective development of cardiorespiratory endurance[2,4,5]. Third, divergence between cross-sectional and longitudinal profiles is usually cultural change (i.e., cohort effects). Data were accumulated from 2011 to 2019, which is long enough for cohort effects to materialize. The variance of between cohort differences in physical fitness was by far the smallest source examined in this study. Fourth, several important covariates (e.g., physical activity, sedentary behaviour, socioeconomic indicators) which affect physical fitness were not assessed but could provide an additional insight into the development of physical fitness. Fifth, physical fitness is highly related to biological maturity and more mature youth outperform less mature youth in physical fitness[40]. Our study indicates the strongest test of an early detection of puberty-related divergence with physical-fitness tests that we are aware of. We conclude that puberty-related development has not started or is not strong enough yet in eight-year old children to be picked up with physical fitness or some of its components. A joint analysis with anthropometric measures (i.e., body mass, body height) and status of biological maturity (e.g., peak height velocity, secondary sex characteristics) will allow a stronger test of the hypothesis that the onset of puberty can already be detected in the ninth year of life.

To sum up, the short-term ontogenetic results of this cross-sectional study revealed test-specific age and sex effects, but no interaction between age and sex despite an abundance of statistical power. According to Ortega et al.[14] physical fitness data of an individual should be compared with reference values of a sex and age-matched similar general population. Such norms are usually only available on annual or semi-annual basis. Our results suggest that physical education teachers, coaches, or researchers can use a proportional adjustment to adequately evaluate physical fitness of prepubertal school-aged children. Furthermore, especially muscle strength / powerUP should be promoted in sport activities for girls in order to reduce the large sex difference between boys and girls. Lastly, physical education teachers, coaches, or researchers should be careful in the interpretation of the ball push test and should consider that it does not measure physical fitness to the same degree than the other tests.

Methods

Sample and study design

This cross-sectional study is part of the EMOTIKON project. The study was mandated and approved by the Ministry of Education, Youth and Sport of the Federal State of Brandenburg, Germany. The Brandenburg School Law[41] requires that parents are comprehensively informed prior to the start of the study. Consent is not needed given that the tests are obligatory for both, children and schools. Physical fitness tests were administered to all third-graders in the state annually between September and November from 2011 to 2019. Physical fitness tests were also administered to 2009 and 2010 cohorts, but later in the school year that is between March and April. Due to the seasonal variation in physical fitness these data were not included. Research was conducted according to the latest Declaration of Helsinki.

We started with data from 144,045 children. Of those, we included only healthy children who had been enrolled within the legal key date of the Federal State of Brandenburg, that is in a given year of school enrolment they were at least 6.00 and at most 6.99 years old on September 30th and, therefore, varied between 8.00 and 8.99 years in the third grade (n = 110,669). In addition to early-entry (n = 2,664), late-entry (n = 30,457) and children without information about birthdate (n = 255), we did not include children with signs of emotional (e.g., autism) and/or physical disorders (e.g., disabilities like infantile cerebral palsy) that were evaluated by the responsible and experienced physical education teacher based on a medical clearance (n = 28). After the first iteration, the LMM-based conditional means of the random effects identified one school as an extreme outlier on several tests across the years and the 171 children of that school were excluded as well. Finally, we applied a +/-3 SD criterion to individual test scores which led to the exclusion of another 2,175 children (2%). This left us with 108,295 children (i.e., 75% of those tested) from 515 different schools.

Physical fitness tests

Physical fitness was assessed with the EMOTIKON test battery (www.uni-potsdam.de/en/emotikon/projekt/methodik for further information on the test protocols). The five tests measured cardiorespiratory endurance (i.e., 6 min run test), coordination (i.e., star run test), speed (i.e., 20-m linear sprint test), power of lower limbs (powerLOW [i.e., standing long jump test]), and power of upper limbs (powerUP [i.e., ball push test]). The EMOTIKON test battery officially includes six tests. Up to 2015 the sixth test was the stand and reach test (flexibility) that was then exchanged against the single-leg balance test (balance). Due to the much smaller of number of scores and their confound with cohort these tests were not included in the analyses. The five tests yielded 525,126 scores from the 108,295 children (i.e., 3% missing test scores).

Qualified physical education teachers of each school administered the tests according to standardized test protocols during the regular physical education classes in the participating schools (www.uni-potsdam.de/en/emotikon/projekt/methodik for further information on the test protocols). Teachers were instructed in a standardized assessment through an advanced training. Tests were always conducted in the morning between 8 and 12 o'clock. Encouragement to achieve the best performance was permitted. Prior to testing, all third-graders performed a standardized warm-up program consisting of different running exercises (e.g., side-steps) and small games (e.g., playing tag).

Cardiorespiratory endurance

Cardiorespiratory endurance was assessed with the 6 min run test. Children had to run as far as they could within six minutes around an official volleyball field (9 x 18 m, every 9 m a pylon/marker was set beside the running court [i.e., six pylons around the field]) at a self-paced velocity. Split time was given every minute. The maximal distance achieved during the six minutes in meters to the nearest nine-meters marker was used as dependent variable in the analysis. The 6 min run test was reliable (test-retest) in children aged 7 to 11 years with an intraclass correlation coefficient (ICC) of 0.92[42]. The 6 min run test correlated at $r = 0.69$ ($p < 0.01$) with $\dot{V}O_{2max}$ assessed via a gas analysis during a progressive treadmill test in children aged 9 to 11 years[43].

Coordination

Coordination under time pressure was tested with the star run test (see Figure 5). Children had to complete a parkour with different movement directions and movement forms (i.e., running forward, running backward, side-steps to the left side, side-steps to the right side). The parkour had to be performed in a given order over a 9 x 9 m star-shaped area where each of the four spikes is marked by a pylon. After starting in the centre of the star, children had to complete the parkour as fast as they could by running in every movement form two times within the given sequence. They had to touch each pylon with the hand. The whole covered distance is 50.912 m. The faster of two test trials was used in the analysis. The shortest time for completing the parkour in seconds to the nearest 1/10 second was measured using a stopwatch and was used as dependent variable in the analysis. The star run test was reliable (test-retest) in 8 to 10 year old children with an ICC of 0.68[44].

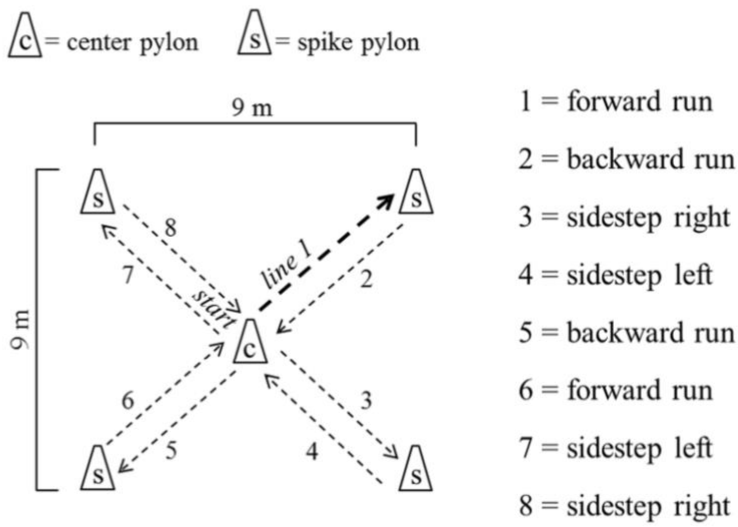

Figure 5. Schematic description of the star run test (adapted from Golle et al.[6]).

Speed

Speed was assessed with the 20-m linear sprint test. After an acoustic signal, the children had to sprint out of a frontal erect posture as fast as they could over a distance of 20 m for two times; the faster test trial was used in the analysis. The shortest time for sprinting the 20 m in seconds to the nearest 1/10 second was measured using a stopwatch and was used as dependent variable in the analysis. The 20-m linear sprint test was reliable (test-retest) in children aged 7 to 11 years with an ICC of 0.90[42].

Power of lower limbs (PowerLOW)

PowerLOW was tested using the standing long jump test. Out of a standing frontal posture the children had to jump as far as they could. The participants had to land with both feet together. They were allowed to swing their arms prior to and during the jump, but after landing the hands were not allowed to touch the floor. The distance in meters to the nearest one centimeter between toes at take-off and heels at landing was determined using a measuring tape; the better of two test trials was used in the analysis. The standing long jump test was reliable (test-retest) in children aged 6 to 12 years with an ICC of 0.94[45].

Power of upper limbs (PowerUP)

PowerUP was assessed through the ball push test. From a standing position the children had to push a 1 kg medicine ball starting in front of the chest with both hands as far as they could for two times; the better of two test trials of longest pushing distance was used in the analysis. The maximal ball push distance in meters to the nearest ten centimeters was determined with a measuring tape and used as dependent variable in the analysis. The ball push test was reliable (test-retest) in children aged 8 to 10 years with an ICC of 0.81[44].

Statistics

Pre- and post-processing of data were carried out in the R environment of statistical computing[46] using the *tidyverse* package[47]. For measures of cardiorespiratory endurance (i.e., 6 min run test), powerLOW (i.e., standing long jump test) and powerUP (i.e., ball push test), higher scores indicated better physical fitness. For measures of coordination (i.e., star run test) and speed (i.e., 20-m linear sprint test), a Box-Cox distributional analyses indicated that a reciprocal transformation brought scores in line with the assumption of a normal distribution[48]. Therefore, we converted scores from seconds to meters/seconds (i.e., pace scores; star run test = 50.912 [m] / time [s]; 20-m linear sprint test = 20 [m] / time [s]). These transformations also had the advantage that a large value was indicative of a good physical fitness for all five measures.

For each test, we determined the ± 3 SD boundary separately for boys and girls. Measurement outside these boundaries were usually implausible (i.e., recording errors) or extreme outliers. They were treated as missing values (3%). Finally, we converted scores within tests (aggregated over boys and girls) to z-scores to facilitate comparison of test, age and sex effects.

Statistical inference was based on a linear mixed model (LMM) estimated with the *MixedModels* package[49] in the *Julia* programming language[50]. The LMM included child (N = 108,295), school (N = 515), and cohort (N = 9) as three random factors; the total number of observations (i.e., max = 5 per child) was 525,126. Fitted model objects were processed with random-effects principal component analysis to obtain loadings of the variance-covariance matrix of the random effects and facilitate its interpretation.

As fixed effects, we specified four sequential-difference contrasts for the five tests: (H1) coordination vs. cardiorespiratory endurance, (H2) speed vs. coordination, (H3) powerLOW vs.

speed, and (H4) powerUP vs. powerLOW. Also included were the effect of age (centered at 8.5 years) as a second-order polynomial trend, the effect of sex (boys – girls), and all interactions between contrasts, age, and sex. Given the large number of observations, children, and schools, we adopted a two-sided z-value > 3.0 as significance criterion for the interpretation of fixed effects.

Child, school, and cohort were included as random factors. With three random factors there was a need for selecting a random-effect structure that included theoretically relevant and reliable variance components (VCs) and correlation parameters (CPs), but was also still supported by the data (i.e., was not overparameterized). Tests varied within children, schools, and cohorts; age and sex varied between children, but within schools and within cohorts. Therefore, in principle, VCs and CPs of linear effects of age and sex could be estimated for schools and cohorts, but not for children.

Parsimonious model selection occurred in two major steps without knowledge or consideration of fixed-effect estimates[51]; details are provided in Supplement A. We started with a model including Grand Mean (varying intercepts) for all three random factors and, given the large numbers of 108,926 children and 515 schools and the small number of nine cohorts, included also test-related VCs and CPs for child and school and age-related and sex-related VCs and CPs for school, but not for cohort. This LMM *m1* was well supported by the data. In the second major step, we increased the complexity of the random-effect structure for cohort by adding test-related VCs (LMM *m2*), then test-related CPs (LMM *m3*), and finally age- and sex-related VCs and CPs (LMM *m4*).

LMM *m4* was not supported by the data (i.e., the fit was singular) and did not significantly improve the goodness of fit over LMM *m3*; delta χ^2 (13) =14.11, p = 0.37. LMM *m3* improved the goodness of fit over LMM *m2* according to the likelihood ratio test, χ^2 (10) = 48.45, $p < 0.001$, but not when the increase in model complexity is penalized according to BIC (i.e., LMM *m2* = 1.27609e6 and LMM *m3* = 1.27617e6). As we had no directed hypotheses relating to test-related CPs for the factor cohort, we stayed with LMM *m2* which represented a very large improvement in goodness of fit relative to LMM *m1*; χ^2 (4) = 1489.57, $p < 0.001$. We also estimated LMM *m2* with two alternative parameterizations that did not change the goodness of fit, but yielded information about CPs between test scores instead of test effects (i.e., contrasts). Finally, we fitted two control LMMs to test the significance of quadratic age trends for fixed effects and the absence of evidence for sex x age interactions separately for each fitness component (i.e., nested within the five levels of the factor test).

Data availability

The datasets generated and analyzed during the current study as well as Julia and R scripts are available in the Open Science Framework (OSF) repository: https://osf.io/2d8rj/

References

1. Malina, R. M., Bouchard, C. & Bar-Or, O. *Growth, Maturation, and Physical Activity*. (Human Kinetics, 2004).
2. Tomkinson, G. R. et al. European normative values for physical fitness in children and adolescents aged 9-17 years: Results from 2 779 165 Eurofit performances representing 30 countries. *Br. J. Sports Med.* **52**, 1445–1456 (2018).
3. De Miguel-Etayo, P. et al. Physical fitness reference standards in European children: The IDEFICS study. *Int. J. Obes*. **38**, S57–S66 (2014).
4. Santos, R. et al. Physical fitness percentiles for Portuguese children and adolescents aged 10-18 years. *J. Sports Sci*. **32**, 1510–1518 (2014).
5. Tambalis, K. D. et al. 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. *Eur. J. Sport Sci.* **16**, 736–746 (2016).
6. Golle, K., Muehlbauer, T., Wick, D. & Granacher, U. Physical fitness percentiles of german children aged 9-12 Years: Findings from a longitudinal study. *PLoS One* **10**, 9–11 (2015).
7. Niessner, C. et al. Representative Percentile Curves of Physical Fitness From Early Childhood to Early Adulthood: The MoMo Study. *Front. Public Heal*. **8**, 1–9 (2020).
8. Stratton, G. & Oliver, J. L. The impact of growth and maturation on physical performance. In *Strength and Conditioning for Young Athletes: Science and Application* (eds Lloyd, R. S. & Oliver, J. L.) 3–20 (Routledge, 2020).
9. Marshall, W. A. & Tanner, J. M. Variations in pattern of pubertal changes in girls. *Arch. Dis. Child.* **44**, 291–303 (1969).
10. Rosenfield, R. L., B.Lipton, R. & Drum, M. L. Thelarche, pubarche, and menarche attainment in children with normal and elevated body mass index. *Pediatrics* **123**, 84–88 (2009).
11. Marshall, W. A. & Tanner, J. M. Variations in pattern of pubertal changes in boys. *Arch. Dis. Child.* **45**, 13–23 (1970).
12. Drenowatz, C., Hinterkörner, F. & Greier, K. Physical Fitness and motor competence in upper Austrian elementary school children — study protocol and preliminary findings of a state-wide fitness testing program. *Front. Sport. Act. Living* **3**, 1–11 (2021).

13. Kanehisa, H., Yata, H., Ikegawa, S. & Fukunaga, T. A cross-sectional study of the size and strength of the lower leg muscles during growth. *Eur. J. Appl. Physiol. Occup. Physiol.* **72**, 150–156 (1995).

14. Ortega, F. B. et al. Physical fitness levels among European adolescents: The HELENA study. *Br. J. Sports Med*. **45**, 20–29 (2011).

15. Dobosz, J., Mayorga-Vega, D. & Viciana, J. Percentile values of physical fitness levels among polish children aged 7 to 19 years – a population-based study. *Cent. Eur. J. Public Health* **23**, 340–351 (2015).

16. Altman, D. G. & Bland, J. M. Absence of evidence is not evidence of absence. *BMJ* **311**, 485 (1995).

17. Overman, W. H. Sex differences in early childhood, adolescence, and adulthood on cognitive tasks that rely on orbital prefrontal cortex. *Brain Cogn*. **55**, 134–147 (2004).

18. Caspersen, C., Powell, K. & Christenson, G. Physical Activity, Exercise, and Physical Fitness: Definitions and Distinctions for Health-Related Research CARL. *Notes Queries* **100**(2), 125–131 (1985).

19. Hattie, J. A. C. *Visible learning: A synthesis of over 800 meta-analyses relating to achievement.* (Routledge, 2009).

20. Fühner, T., Kliegl, R., Arntz, F., Kriemler, S. & Granacher, U. An Update on Secular Trends in Physical Fitness of Children and Adolescents from 1972 to 2015: A Systematic Review. *Sport. Med*. **51**, 303–320 (2021).

21. Tomkinson, G. R., Lang, J. J. & Tremblay, M. S. Temporal trends in the cardiorespiratory fitness of children and adolescents representing 19 high-income and upper middle-income countries between 1981 and 2014. *Br. J. Sports Med*. **53**, 478–486 (2019).

22. Spengler, S., Rabel, M., Kuritz, A. M. & Mess, F. Trends in motor performance of first graders: A comparison of cohorts from 2006 to 2015. *Front. Pediatr*. **5**, 1–7 (2017).

23. Woll, A., Kurth, B. M., Opper, E., Worth, A. & Bös, K. The 'Motorik-Modul' (MoMo): Physical fitness and physical activity in German children and adolescents. *Eur. J. Pediatr*. **170**, 1129–1142 (2011).

24. Catley, M. J. & Tomkinson, G. R. Normative health-related fitness values for children: Analysis of 85347 test results on 9-17-year-old Australians since 1985. *Br. J. Sports Med.* **47**, 98–108 (2013).

25. Roriz de Oliveira, M. S., Seabra, A., Freitas, D., Eisenmann, J. C. & Maia, J. Physical fitness percentile charts for children aged 6-10 from Portugal. *J. Sports Med. Phys. Fitness* **54**, 780–792 (2014).

26. Ramos-Sepúlveda, J. A., Ramírez-Vélez, R., Correa-Bautista, J. E., Izquierdo, M. & García-Hermoso, A. Physical fitness and anthropometric normative values among Colombian-Indian schoolchildren. *BMC Public Health* **16**, 1–15 (2016).

27. Beunen, G. & Thomis, M. Muscular strength development in children and adolescents. *Pediatr. Exerc. Sci.* **12**, 174–197 (2000).

28. Round, J. M., Jones, D. A., Honour, J. W. & Nevill, A. M. Hormonal factors in the development of differences in strength between boys and girls during adolescence: A longitudinal study. *Ann. Hum. Biol.* **26**, 49–62 (1999).

29. Armstrong, N. & Welsman, J. Aerobic fitness: What are we measuring? *Med. Sport Sci.* **50**, 5–25 (2007).

30. Haywood, K. M. & Getchell, N. *Life span motor development*. (Human Kinetics, 2009). doi:10.1016/0167-9457(87)90023-6.

31. Ardila, A., Rosselli, M., Matute, E. & Inozemtseva, O. Gender Differences in Cognitive Development. *Dev. Psychol.* **47**, 984–990 (2011).

32. Ludyga, S., Gerber, M., Pühse, U., Looser, V. N. & Kamijo, K. Systematic review and meta-analysis investigating moderators of long-term effects of exercise on cognition in healthy individuals. *Nat. Hum. Behav.* **4**, 603–612 (2020).

33. Schmidt, M. et al. Disentangling the relationship between children's motor ability, executive function and academic achievement. *PLoS One* **12**, (2017).

34. Koutsandréou, F., Wegner, M., Niemann, C. & Budde, H. Effects of motor versus cardiovascular exercise training on children's working memory. *Med. Sci. Sports Exerc.* **48**, 1144–1152 (2016).

35. Viru, A. et al. Critical Periods in the Development of Performance Capacity During Childhood and Adolescence. *Eur. J. Phys. Educ.* **4**, 75–119 (1999).

36. Viru, A. et al. Age periods of accelerated improvemnt of muscle strength, power, speed and endurance in the age interval 6-18 yeras. *Biol. Sport* **15**, 211–227 (1998).

37. Ducher, G. et al. Overweight children have a greater proportion of fat mass relative to muscle mass in the upper limbs than in the lower limbs: Implications for bone strength at the distal forearm. *Am. J. Clin. Nutr.* **90**, 1104–1111 (2009).

38. García-Hermoso, A., Ramírez-Vélez, R., Lubans, D. R. & Izquierdo, M. Effects of physical education interventions on cognition and academic performance outcomes in children and adolescents: a systematic review and meta-analysis. *Br. J. Sports Med*. bjsports-2021-104112 (2021) doi:10.1136/bjsports-2021-104112.

39. Vrbik, I. et al. The Influence of Familiarization on Physical Fitness Test Results in Primary School-Aged Children. *Pediatr. Exerc. Sci*. **29**, 278–284 (2017).

40. Jones, M. A., Hitchen, P. J. & Stratton, G. The importance of considering biological maturity when assessing physical fitness measures in girls and boys aged 10 to 16 years. *Ann. Hum. Biol*. **27**, 57–65 (2000).

41. Ministerium für Bildung Jugend und Sport. Gesetz über die Schulen im Land Brandenburg. https://bravors.brandenburg.de/gesetze/bbgschulg (2018).

42. Bös, K. *Deutscher Motorik-Test 6-18*. (Czwalina, 2009).

43. von Haaren, B., Härtel, S., Seidel, I., Schlenker, L. & Bös, K. Validity of a 6-min Endurance Run and a 20-m Shuttle Run in 9-to 11-Year Old Children. *Dtsch. Z. Sportmed*. **62**, 351–355 (2011).

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

45. Fernandez-Santos, J. R., Ruiz, J. R., Cohen, D. D., Gonzalez-Montesinos, J. L. & Castro-Pinero, J. Reliability and validity of tests to assess lower-body muscular power in children. *J. Strength Cond. Res.* **29**, 2277–2285 (2015).

46. Team, R. C. R a language and environment for statistical computing. https://www.r-project.org/ (2020).

47. Wickham, H. et al. Welcome to the Tidyverse. *J. Open Source Softw*. **4**, 1686 (2019).

48. Box, G. E. P. & Cox, D. R. An analysis of transformations. *J. R. Stat. Soc. Ser. B* **26**, 211–243 (1964).

49. Bates, D. et al. JuliaStats/MixedModels.jl:v3.4.0. vol. 1 20071787–20071787 https://doi.org/10.5281/zenodo.4589101 (2021).

50. Bezanson, J., Edelman, A., Karpinski, S. & Shah, V. B. Julia: A fresh approach to numerical computing. *SIAM Rev.* **59**, 65–98 (2017).

51. Bates, D., Kliegl, R., Vasishth, S. & Baayen, H. Parsimonious Mixed Models. ArXiv:1506.04967 http://arxiv.org/abs/1506.04967 (2015).

Funding

Open Access funding enabled and organized by Projekt DEAL. The study was commissioned and funded by the Ministry of Education, Youth, and Sport of the Federal State of Brandenburg, Germany. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript. Reinhold Kliegl was supported by the Center for Interdisciplinary Research, Bielefeld (ZiF)/Cooperation Group "Statistical models for psychological and linguistic data".

Acknowledgement

The authors acknowledge extensions of MixedModels.jl software by Phillip Alday, Douglas Bates, and Dave Kleinschmidt that were necessary for fitting the linear mixed models. Ralph Krampe, Christian Puta, Daniel Schad, and three reviewers provided helpful comments on the manuscript. The authors acknowledge the support of the Deutsche Forschungsgemeinschaft and Open Access Publishing Fund of University of Potsdam.

Author contributions

TF, KG, and UG contributed to conception and design; TF and KG organized data collection; RK and TF carried out data analysis; TF and RK wrote the first draft of the manuscript and all authors were involved in iterative revisions; all authors provided final approval of the version to be published and agreed to be accountable for all aspects of the work.

Competing Interests Statement

Thea Fühner, Urs Granacher, Kathleen Golle, and Reinhold Kliegl declare that they have no conflicts of interest relevant to the content of this cross-sectional study.

Publication III

EFFECT OF TIMING OF SCHOOL ENROLLMENT ON PHYSICAL FITNESS IN THIRD GRADERS

Thea Führer[1], Urs Granacher[1], Kathleen Golle[1], Reinhold Kliegl[1]

[1] University of Potsdam, Research Focus Cognition Sciences, Division of Training and Movement Science, Potsdam, Germany

Reference

Führer, T., Granacher, U., Golle, K., & Kliegl, R. (2022). Effect of timing of school enrollment on physical fitness in third graders. *Scientific Reports, 12,* 7801. https://doi.org/10.1038/s41598-022-11710-x

The final publication is available at:

https://www.nature.com/articles/s41598-022-11710-x

Effect of timing of school enrollment on physical fitness in third graders

Thea Fühner[1*], Urs Granacher[1], Kathleen Golle[1], Reinhold Kliegl[1]

[1]University of Potsdam, Faculty of Human Sciences, Division of Training and Movement Sciences, Am Neuen Palais 10, Building 12, 14469 Potsdam, Germany

[*]Corresponding author and contact details

Thea Fühner

University of Potsdam

Faculty of Human Sciences

Division of Training and Movement Sciences

Am Neuen Palais 10, Building 12

14469 Potsdam

Germany

Tel: +49-331-977 1672

E-mail address: fuehner@uni-potsdam.de

ORCID ID: 0000-0002-1120-0986

Abstract

Timing of initial school enrollment may vary considerably for various reasons such as early or delayed enrollment, skipped or repeated school classes. Accordingly, the age range within school grades includes older- (OTK) and younger-than-keyage (YTK) children. Hardly any information is available on the impact of timing of school enrollment on physical fitness. There is evidence from a related research topic showing large differences in academic performance between OTK and YTK children versus keyage children. Thus, the aim of this study was to compare physical fitness of OTK (N = 26,540) and YTK (N = 2,586) children versus keyage children (N = 108,295) in a representative sample of German third graders. Physical fitness tests comprised cardiorespiratory endurance, coordination, speed, lower, and upper limbs muscle power. Predictions of physical fitness performance for YTK and OTK children were estimated using data from keyage children by taking age, sex, school, and assessment year into account. Data were annually recorded between 2011 and 2019. The difference between observed and predicted z-scores yielded a delta z-score that was used as a dependent variable in the linear mixed models. Findings indicate that OTK children showed poorer performance compared to keyage children, especially in coordination, and that YTK children outperformed keyage children, especially in coordination. Teachers should be aware that OTK children show poorer physical fitness performance compared to keyage children.

Introduction

The importance of physical fitness for children's health is undisputed[1]. According to Caspersen et al.[2], physical fitness can be categorized as health- (e.g., cardiorespiratory endurance, muscular endurance, muscular strength, body composition, and flexibility) or skill-related fitness (e.g., agility, balance, coordination, speed, [muscle] power, and reaction time). There is evidence from original research[3], systematic reviews, and meta-analyses[4,5] that cardiorespiratory endurance and muscular strength are positively associated with markers of physical health (e.g., body mass index, waist circumference, skinfold thickness, cardiovascular disease risk score) in youth. Accordingly, it is important to regularly monitor and evaluate children's physical fitness to identify potential deficits in physical fitness as early as possible. Recent studies on global secular trends in youth physical fitness indicated physical fitness declines particularly for measures of cardiorespiratory endurance. This trend additionally emphasizes the relevance of physical fitness testing[6,7].

Physical fitness tests represent an easy-to-administer, reliable, and valid means to assess and evaluate children's physical fitness in large scale studies conducted in sport clubs or schools[8]. Several studies from around the globe[8–17] showed developmental increases in physical fitness from childhood to adolescence[8,10–12,15,17]. Irrespective of age, boys outperform girls in most components of physical fitness[9–17], except for balance[10,17] and flexibility[9–15,17].

The available studies on physical fitness development have been conducted in youth aged [5-18] years. In these studies, children and adolescents were matched into one-year age groups. This age grouping system is also evident in many settings of children's everyday life. For instance, children are matched in one-year age teams within sport clubs or in grades within schools. However, this age grouping system is not without limitations because of differences in relative age depending on the specific cut-off date under consideration. For schools in general, the cut-off date of initial school enrollment is specific to the country under investigation. For instance, in the Federal State of Brandenburg, Germany the official and initial school enrollment date is September 30th. Accordingly, children are enrolled to school (i.e., first grade) if they are aged between 6 years and 0 months and 6 years and 11 months on September 30th of the respective year (i.e., keyage children in first grade). Thus, children who are born on September 30th or slightly later are at the extreme end, i.e., almost one year older than their classmates who are born in August. These differences in the birthdate may have an impact on anthropometrics (e.g., body height, body mass) and physical fitness (e.g., muscular strength, power, cardiorespiratory endurance, or speed)[18] because physical fitness performance increases with age from childhood to adolescence[8,10–12,15,17]. Thus, within a one-year age-group, the relatively older children (i.e., born near the cut-off date) may outperform their relative younger classmates (i.e., born later to the cut-off date) because of their relatively older age[18,19]. In fact, a previous study conducted with keyage third graders (i.e., children aged 8 years and 0 months to 8 years and 11 months) has shown that physical fitness increased linearly with chronological age[20]. Furthermore, even within the single ninth year of life, the relatively older children (i.e., aged 8 years and 6 months to 8 years and 11 months) significantly outperformed the younger children (i.e., aged 8 years and 0 months to 8 years and 5 months) in physical fitness[20].

Within one school-grade, there are keyage children as well as younger- (YTK) or older-than-keyage (OTK) children. This is due to early or late school enrollment, skipping or repetition of a school year. With reference to our data, age ranged from 5 years and 11 months to 14 years and 5 months for our study sample that included YTK and OTK children. Given that there are already large differences in physical fitness within the group of keyage children[20], the question arises as to physical fitness performance of YTK and OTK children. To the authors' knowledge there is hardly any information available in the literature on differences in physical fitness of YTK and OTK children versus keyage children. A major goal of physical education is to create a learning setting for each child according to his / her individual needs to ensure a holistic development. Thus, findings on physical fitness performance of YTK and OTK children provide valuable information to promote physical fitness according to the child´s individual needs. For instance, children who show delayed physical fitness development should receive additional health and fitness programs to compensate their deficits in physical fitness. Furthermore, given that grading systems are only available for keyage children, findings of this study can be used to individually grade physical fitness according to age, sex, and timing of school enrollment.

Information from a related research topic shows large differences in academic performance between OTK and YTK children versus keyage children[21–23]. For instance, in a study including 1,144 German primary school children, Urschitz et al.[23] reported that especially OTK children aged > 9 years compared with keyage children showed poor academic performance in terms of grades in mathematics, science, reading, spelling, and handwriting. In a study including 3,684 Australian high school students aged 14 years, Martin[22] reported that YTK children scored significantly better in academic performance (i.e., performance in literacy and numeracy) than keyage children. However, as already mentioned this has not yet been examined for physical fitness. Therefore, the aim of this cross-sectional study was to compare physical fitness of OTK and YTK children versus keyage children in a sample of German primary school children taking age, sex, school, and assessment year into account. With reference to the relevant school-based studies on differences in academic performance of OTK and YTK children versus keyage children[21–23], we hypothesized that OTK children show poorer and YTK children better physical fitness performance compared with keyage children.

Methods

Experimental approach

This cross-sectional study is part of the ongoing EMOTIKON research project (www.uni-potsdam.de/en/emotikon). Physical fitness tests were conducted every year between September and November starting in 2011. Physical fitness tests were also administered in 2009 and 2010, but later in the school year that is between March and April. Due to the seasonal variation in physical fitness these data were not included.

Population

Since 2009, all third graders living in the Federal State of Brandenburg, Germany were tested annually for their physical fitness. This cross-sectional study was mandated and approved by the Ministry of Education, Youth and Sport of the Federal State of Brandenburg, Germany. The Brandenburg School Law requires that parents are comprehensively informed prior to the start of the study. Consent is not needed given that the tests are obligatory for both, children and schools[24]. None of the authors included in the author list had access to personally identifiable information on the children. The authors received the data absolutely anonymized from the Ministry of Education, Youth and Sport of the Federal State of Brandenburg, Germany. Research was conducted according to the latest Declaration of Helsinki[25].

To compare physical fitness development of YTK and OTK children with that of keyage children, we used physical fitness data recorded between 2011 and 2019.

- 2,586 YTK children aged 7 years and 0 months to 7 years and 11 months
- 108,296 keyage children aged 8 years and 0 months to 8 years and 11 months
- 26,540 OTK children aged 9 years and 0 months to 9 years and 11 months

Selection into keyage, OTK, and YTK groups was strictly based on children's birthdate relative to the legal date for school enrollment (i.e., September 30th in the Federal State of Brandenburg for all assessment years). Thus, on September 30th, keyage third graders ranged between 8 years and 0 months to 8 years and 11 months. YTK children were younger, and OTK children were older.

The selection of keyage children has been described in a previous publication of our research group[20]. Data from an earlier study were used as a reference for OTK and YTK children. Initially, 30,253 OTK children were included in the data base: 2,842 were excluded due to age. The excluded third-graders ranged from 10 years and 1 month to 14 years and 5 months. Another 27 students were excluded due to adverse health events as reported by the responsible teacher (e.g., physical disability, autism spectrum). Finally, 844 students were considered outliers and outside +/- 3 SD of their group x sex x test cell. Finally, 26,540 OTK children were included in the analyses (88.7%). For YTK children, initially 2,654 YTK were eligible to be included in the data base. From this initial sample, 28 were excluded due to age because they ranged from 5 years and 11 months to 6 years and 11 months. Moreover, 40 children were considered outliers and beyond +/- 3 SD in their group x sex x test cell. Finally, 2,586 YTK children were included in the analyses (97.4%).

Physical fitness tests

Physical fitness was assessed using the specific EMOTIKON test battery[20]. These tests evaluated cardiorespiratory endurance (i.e., 6-minute-run test), coordination (i.e., star-run test), speed (i.e., 20-m linear sprint test), lower (powerLOW [i.e., standing long jump test]), and upper limbs muscle power (i.e., powerUP [ball-push test]). The EMOTIKON test battery officially includes six tests. In 2016, the assessment of flexibility (i.e., stand-and-reach test) was stopped and the assessment of balance (i.e., single-leg balance test with eyes closed) was included[26]. Due to the much smaller number of scores and their confound with assessment year, these two tests were not included in the analyses.

Physical fitness tests were administered by qualified physical education teachers and conducted during the regular physical education classes. All physical education teachers received standardized test instructions for the assessment (www.uni-potsdam.de/en/emotikon/projekt/methodik - for further information on the test protocols). Furthermore, all physical education teachers participated in advanced training programs about standardized physical fitness assessment. Tests were always conducted in the morning between 8 and 12 am. Prior to testing, all third-graders performed a standardized warm-up program consisting of different running exercises (e.g., side-steps) and small games (e.g., playing tag).

Cardiorespiratory endurance

Cardiorespiratory endurance was assessed using the 6-minute-run test. Participating children had to run the furthest distance during the six minutes test time around a volleyball field (54 m) at a self-paced velocity. The test instructor provided split times every minute. After the six minutes, maximal distance covered in meters to the nearest nine meters was recorded and used as dependent variable. High test-retest reliability was reported for the 6-minute-run test with an intraclass correlation coefficient (ICC) of 0.92 in children aged 7 to 11 years[27].

Coordination

Coordination under time pressure was evaluated using the star-run test. During the star-run test, the participating children had to complete a parkour with different running techniques (i.e., running forwards, running backwards, side-steps) as quickly as possible. The star shaped parkour (9 m x 9 m) consisted of four spikes. Each spike and the center of the star were marked with a pylon. The participants started in the middle of the star. First, they had to run forward to the first pylon and backward to the middle. Next, they had to do side-steps to the second pylon on the right side and side-steps back to the middle. Then, they had to run backward to the third pylon and forward to the middle. Finally, they had to do side-steps to the fourth pylon on the left side and side-steps back to the middle. The participants had to touch each pylon within the parkour with the hand. The whole covered distance was 50.912 m. Time for test completion in seconds to the nearest 1/10 second was taken using a stopwatch and used as dependent variable in the analysis. The participants had two test trials of which the best test trial in terms of time until test completion was kept for analysis. The star run test was reliable (test-retest) for children aged 8 to 10 years with an ICC of 0.68[28].

Speed

Speed was assessed using the 20-m linear sprint test. The participating children started from a standing position with one foot right behind the starting line. After an acoustic signal, they had to sprint as fast as possible over a distance of 20 m. Time for test completion in seconds to the nearest 1/10 second was taken using a stopwatch and used as dependent variable in the analysis. The participants had two test trials of which the best trial was taken for further analysis in terms of the time until test completion. Test-retest reliability has been reported to be high for children aged 7 to 11 years with an ICC of 0.90[27].

Lower limbs muscle power (PowerLOW)

PowerLOW was assessed through the standing long jump test. The participating children had to jump as far as possible from a frontal position. Arm swing prior to and during the jump was allowed. Jump distance in centimeters between the starting line and heel of the posterior foot was recorded to the nearest one centimeter using a measuring tape. The participants had two test trials of which the best trial in terms of the longest jump distance was taken for further analysis. The standing long jump test showed high test-retest reliability for children aged 6 to 12 years with an ICC of 0.94[29].

Upper limbs muscle power (PowerUP)

Power up was evaluated with the ball-push test. From a standing position, the participating children had to push a 1 kg medicine ball that was held tight right in front of the chest. The participants had to push the ball at maximal effort with both hands. The pushing distance in meters was recorded to the nearest ten centimeters using a measuring tape. The participants had two test trials of which the best test trial in terms of the longest pushing distance was taken for further analysis. The ball-push test was reliable (test-retest) for children aged 8 to 10 years with an ICC of 0.81[28].

Statistics

Pre- and post-processing of data were carried out in the R environment of statistical computing[30] using the *tidyverse* package[31]. For statistical inference we relied on Linear Mixed Model analyses (LMM) with the *MixedModels* package[32] in the *Julia* programming language (v 1.7.1)[33].

For measures of cardiorespiratory endurance (i.e., 6 min run test), powerLOW (i.e., standing long jump test) and powerUP (i.e., ball push test), higher scores indicate better physical fitness. For measures of coordination (i.e., star run test) and speed (i.e., 20-m linear sprint test), a Box-Cox distributional analyses indicated that a reciprocal transformation brought scores in line with the assumption of a normal distribution[34]. Therefore, we converted scores from seconds to meters/seconds (i.e., pace scores; star run test = 50.912 [m] / time [s]; 20-m linear sprint test = 20 [m] / time [s]). These transformations had the advantage that a large value was indicative of good physical fitness for all five tests. Finally, z-scores were computed in two stages. In the first stage, we calculated z-scores within the test (i.e., 6-minute-run test, star-run test, 20-m linear sprint test, standing long jump test, ball-push test) x sex (male, female) x group (YTK, OTK) cells and removed observations exceeding +/- 3 SDs (i.e., outliers). This is in accordance with a previous publication from the same research group[20]. In the second stage, we used means and SDs of the five fitness tests for keyage children from a previous study[20] and computed the respective z-scores that were included in figures 1 and 2.

To compare YTK and OTK children`s development of physical fitness with that of keyage children, we predicted the physical fitness performance for ages 7 years and 0 months to 7 years and 11 months and 9 years and 0 months to 9 years and 11 months using LMM parameter estimates of the 108,295 keyage children (i.e., grey lines in figure 1), reported in Fühner et al[20]. Through this predication analyses we received the information about physical fitness performance of keyage children at the ages 7 years and 0 months to 7 years and 11 months and 9 years and 0 months and 9 years and 11 months. The model parameters comprised fixed effects for age, tests, sex, and their interactions, variance components (VCs) and correlation parameters (CPs) for GM and four test contrasts for the random factor child, VCs and CPs for GM, four test contrasts, sex, and age for the random factor school, and VCs for test and, age for the random factor assessment year. Details about model specification for these predictions are provided in Fühner et al.[20] and in script: fggk22_lmm_pred.jl in the repository.

Through physical fitness testing in EMOTIKON, we obtained the actual physical fitness status of YTK children aged 7 years and 0 months to 7 years and 11 months and OTK children aged

9 years and 0 months to 9 years and 11 months. Please note that the classification of children into YTK, keyage, or OTK groups is based solely on children's birthdate whereas children's age is the difference between the date of test and their birthdate. Therefore, some YTK children were slightly older than 8 years and some OTK children were slightly younger than 9 years at the time of testing. Results did not change if non-keyage children aged between 8 to 9 years were excluded.

The difference between observed (i.e., obtained through physical fitness testing) and predicted z-scores (i.e., predicted data from keyage children [grey lines in figure 1]) yielded a delta z-score that was used as dependent variable in the following LMMs to compare physical fitness development of YTK and OTK children (i.e., obtained scores through physical fitness testing) with that of keyage children (i.e., predicted data).

We analyzed the data with separate LMMs for OTK and YTK children. The fixed effects included in the starting LMM were similar to the one reported by Fühner et al.[20]. Specifically, there were four sequential-difference fixed-effect contrasts for the five tests: (H1) coordination vs. cardiorespiratory endurance, (H2) speed vs. coordination, (H3) powerLOW vs. speed, and (H4) powerUP vs. powerLOW. We additionally included the effect of age (centered at 8 years and 6 months) as a second-order polynomial trend, the effect of sex (boys – girls), and all interactions between contrasts, age, and sex. We used a two-sided z-value > 2.0 as significance criterion for the interpretation of fixed effects.

The random effect structure included VCs and CPs of the delta z-scores for the five tests related to grouping (random) factors of child, school, and assessment year. Tests varied within children, schools, and assessment years; age and sex varied between children, but within schools and within assessment years. Therefore, in principle, VCs and CPs also include effects of age and sex for the factors school and assessment year.

LMM for older-than-keyage (OTK) children

The initial LMM included child (N = 26,540), school (N = 513), and assessment year (N = 9) as three random factors; the total number of observations (i.e., max = 5 per child) was 128,198. With three random factors, there was a need for selecting a random-effect structure that included theoretically relevant and reliable VCs and CPs but was also still supported by the data (i.e., was not overparameterized).

Parsimonious model selection occurred in two major steps without knowledge or consideration of fixed-effect estimates[35]; details are provided in script: fggk22_lmm_otk.jl in the repository. The random-effect structure of the parsimonious LMM of delta z-scores was expected to be simpler than the one for the LMM of Fühner et al.[20] because the much smaller number of children and, importantly, because most of the school- and assessment-year-related random effects as well as fixed effect of age and sex were included in the predicted z-scores. We started with a model estimating VCs and CPs between delta z-scores of the five tests for children and VCs of delta z-scores for the five tests, age, and sex for school, and only varying intercept (GM) for assessment year. This LMM was well supported by the data. Increasing the complexity of the random-effect structure by adding CPs for school or adding VCs for assessment year did not improve the goodness of fit. Moreover, the school-related VC for sex and high-order fixed-effect interactions between test, age, and sex could be removed without loss of goodness of fit. As in Fühner et al.[20], we also estimated the final model with an alternative post-hoc LMM parameterization to test main fixed effects of sex and age separately for each fitness test (i.e., we specified sex and age as nested within the five levels of the factor test).

LMM for younger-than-keyage (YTK) children

The LMM included child (N = 2,586), school (N = 437), and assessment year (N = 9) as three random factors; the total number of observations (i.e., max = 5 per child) was 12,590. In the model selection process, we followed the model of OTK described above.

Parsimonious model selection occurred without knowledge or consideration of fixed-effect estimates[35]; details are provided in script: fggk22_lmm_ytk.jl in the repository. First, we applied the LMM of OTK to the data of YTK. This model was not supported by the data (i.e., overparameterized) because of the relatively small sample size of YTK (N = 2,586) compared to OTK (N = 26,540). Indeed, the data supported only a LMM with a strongly reduced complexity, comprising (a) fixed effects on delta z-scores for the four contrasts of test, (b) VCs for the five delta z-scores for school and child, and (c) CPs for the five delta z-scores of child. Thus, there was no statistical support for fixed or random effects of age and sex for YTK children relating to delta z-scores.

Results

Table 1 summarizes descriptive statistics for the three subsamples of third-graders. Statistics about keyage children refer to the sample reported in Fühner et al.[20]. Statistics about YTK and OTK children refer to the samples of this study.

Table 1 Descriptive statistics for younger-than-keyage, keyage, and older-than-keyage children

Sample	Physical Fitness Component	Sex	N Schools	N Child	Mean Age [years]	SD Age [years]	Mean score	SD score	Mean *delta*	SD *delta*
Keyage	Endurance [m]	Boys	513	51116	8.56	0.28	1041.38	154.03	0	0.4
Keyage	Endurance [m]	Girls	511	52821	8.55	0.28	967.72	132.50	0	0.3
Keyage	Coordination [m/s]	Boys	512	51023	8.56	0.28	2.08	0.30	0	0.4
Keyage	Coordination [m/s]	Girls	510	52886	8.55	0.28	2.01	0.27	0	0.3
Keyage	Speed [m/s]	Boys	513	51700	8.56	0.28	4.58	0.42	0	0.4
Keyage	Speed [m/s]	Girls	512	53259	8.55	0.28	4.45	0.39	0	0.4
Keyage	PowerLOW [cm]	Boys	513	52141	8.56	0.28	129.41	19.53	0	0.4
Keyage	PowerLOW [cm]	Girls	509	53856	8.55	0.28	122.00	18.44	0	0.4
Keyage	PowerUP [m]	Boys	514	52254	8.56	0.28	3.99	0.70	0	0.4
keyage	PowerUP [m]	Girls	512	54070	8.55	0.28	3.50	0.63	0	0.3
OTK	Endurance [m]	Boys	511	14870	9.35	0.25	1017.86	166.33	-0.18	1.1
OTK	Endurance [m]	Girls	499	10519	9.35	0.26	950.97	140.53	-0.14	1.0
OTK	Coordination [m/s]	Boys	509	14808	9.35	0.25	2.06	0.31	-0.28	1.1
OTK	Coordination [m/s]	Girls	502	10542	9.35	0.26	1.99	0.29	-0.30	1.1
OTK	Speed [m/s]	Boys	511	15010	9.36	0.25	4.58	0.44	-0.17	1.1
OTK	Speed [m/s]	Girls	503	10644	9.35	0.26	4.44	0.41	-0.19	1.1
OTK	PowerLOW [cm]	Boys	511	15137	9.35	0.26	127.83	21.08	-0.23	1.2
OTK	PowerLOW [cm]	Girls	502	10699	9.35	0.26	119.42	19.45	-0.29	1.1
OTK	PowerUP [m]	Boys	511	15236	9.36	0.25	4.13	0.75	-0.22	1.1
OTK	PowerUP [m]	Girls	503	10733	9.35	0.26	3.62	0.67	-0.23	1.0
YTK	Endurance [m]	Boys	350	1087	7.85	0.19	1042.92	149.54	0.036	1.0
YTK	Endurance [m]	Girls	384	1408	7.88	0.18	973.11	132.88	0.035	1.0
YTK	Coordination [m/s]	Boys	350	1091	7.85	0.19	2.05	0.29	0.10	1.1
YTK	Coordination [m/s]	Girls	382	1397	7.87	0.18	1.99	0.26	0.10	1.0
YTK	Speed [m/s]	Boys	349	1097	7.85	0.19	4.51	0.40	0.035	1.1
YTK	Speed [m/s]	Girls	385	1423	7.87	0.18	4.42	0.39	0.070	1.0
YTK	PowerLOW [cm]	Boys	350	1112	7.85	0.19	128.54	18.48	0.082	1.1
YTK	PowerLOW [cm]	Girls	384	1433	7.87	0.18	121.77	18.06	0.078	1.0
YTK	PowerUP [m]	Boys	348	1111	7.85	0.19	3.79	0.70	0.12	1.0
YTK	PowerUP [m]	Girls	384	1431	7.88	0.18	3.33	0.61	0.14	0.9

N = sample size, SD = standard deviation, *delta* = difference between observed (i.e., obtained through physical fitness testing) and predicted z-scores (i.e., predicted data from keyage children [grey lines in figure 1]), Endurance = cardiorespiratory endurance (i.e., 6-minute-run test), Coordination = star-run test, Speed = 20-m linear sprint test, PowerLOW = lower limbs muscle power (i.e., standing long jump test), PowerUP = upper limbs muscle power (i.e., ball-push test), OTK = older-than-keyage children (i.e., aggregated over 9 years and 0 months to 9 years and 11 months), YTK = younger-than-keyage children (i.e., aggregated over 7 years and 0 months to 7 years and 11 months). Coordination and speed times were converted from seconds to meters/seconds (i.e., pace scores; star-run test = 50.912 [m] / time [s]; 20-m linear sprint test = 20 [m] / time [s]). These transformations have the advantage that a large value is indicative of better physical fitness.

Figure 1 displays the observed (points) and predicted (lines) physical fitness development for YTK boys and girls aged 7 years and 0 months to 7 years and 11 months and OTK boys and girls aged 9 years and 0 months to 9 years and 11 months. The predicted z-scores for keyage children aged 8 years and 0 months to 8 years and 11 months are located on the predicted lines. There is a slight overlap between groups at 8- and 9-year boundaries due to birthdate determining the classification of children into keyage groups and age being measured as the difference between age at test and birthdate.

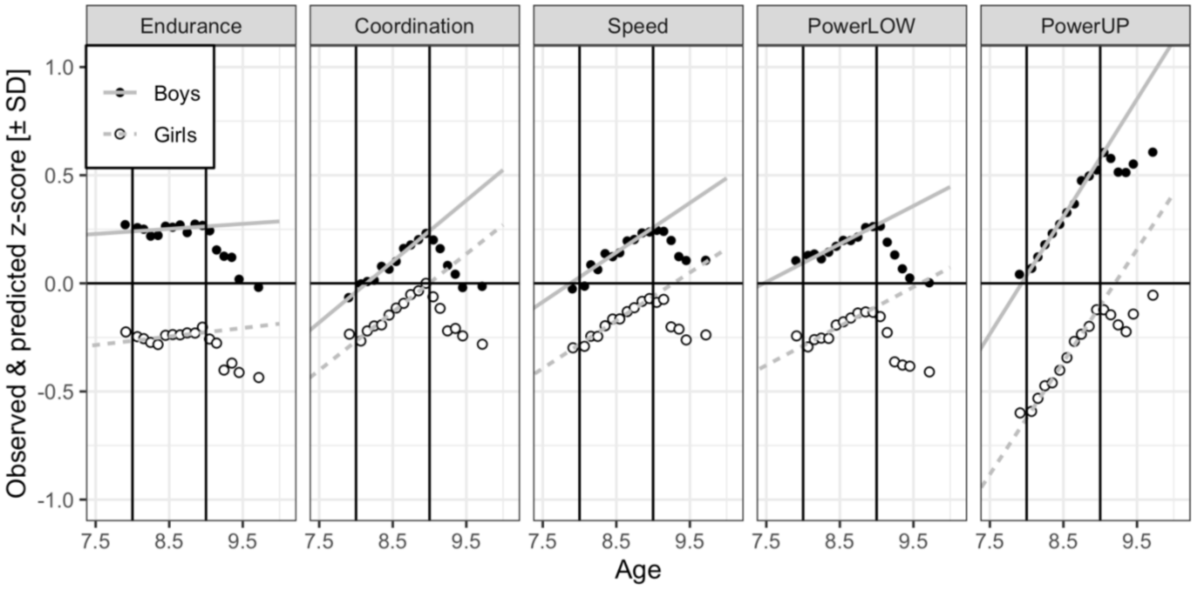

Figure 1. Observed z-scores for physical fitness development for boys (closed circles) and girls (open circles) aged 7.00 to 10.0 years. The lines represent the predicted z-scores for physical fitness development for boys (grey line) and girls (dashed grey line). Data were z-transformed. Endurance = cardiorespiratory endurance (i.e., 6-minute-run test), Coordination = star-run test, Speed = 20-m linear sprint test, PowerLOW = lower limbs muscle power (i.e., standing long jump test), PowerUP = upper limbs muscle power (i.e., ball-push test). Note that *delta z-scores* for younger-than-keyage boys and girls were aggregated over 7.00 to 7.99 years and that *delta z-scores* for older-than-keyage boys and girls were aggregated over 9.50 to 9.99 years. Points are binned observed child means. Coordination and speed times were converted from seconds to meters/seconds (i.e., pace scores; star-run test = 50.912 [m] / time [s]; 20-m linear sprint test = 20 [m] / time [s]). These transformations have the advantage that a large value is indicative of better physical fitness and that they remove skew in the distributions.

Figure 2 displays the delta z-scores between observed and predicted physical fitness development for YTK boys and girls aged 7 years and 0 months to 7 years and 11 months and OTK boys and girls aged 9 years and 0 months to 9 years and 11 months. The delta z-scores for keyage children aged 8 years and 0 months to 8 years and 11 months are represented in the horizontal zero line. The z-scores for OTK and YTK children will be described in the next sections.

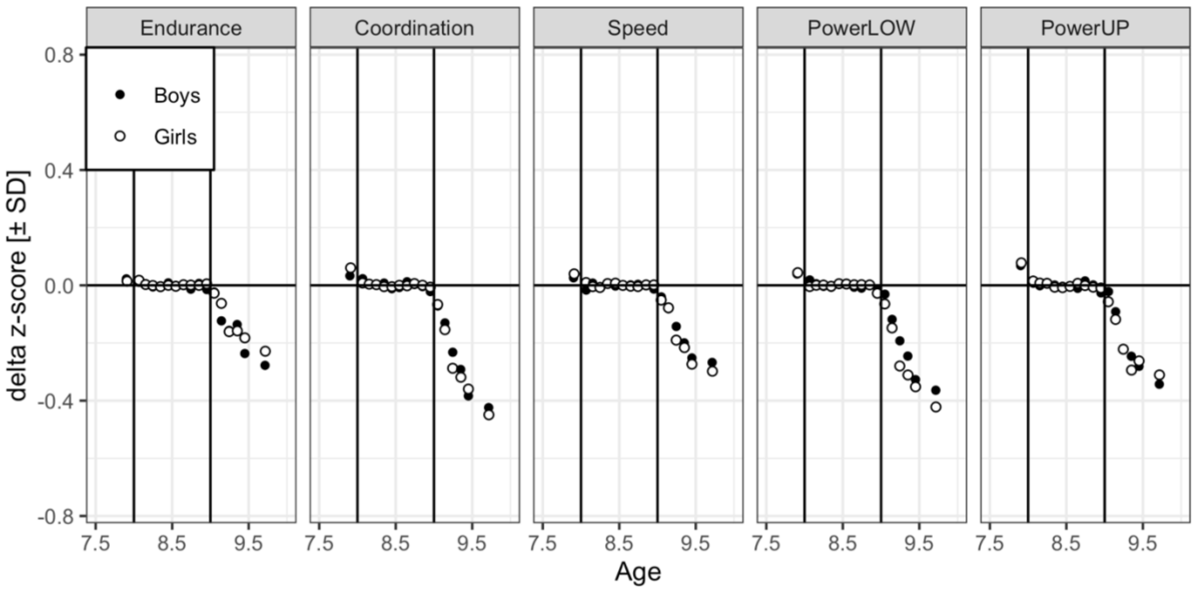

Figure 2. *Delta z-score* between observed and predicted physical fitness development for boys (closed circles) and girls (open circles) aged 7.00 to 10.0 years. Data were z-transformed. Endurance = cardiorespiratory endurance (i.e., 6-minute-run test), Coordination = star-run test, Speed = 20-m linear sprint test, PowerLOW = lower limbs muscle power (i.e., standing long jump test), PowerUP = upper limbs muscle power (i.e., ball-push test). Note that *delta* z-scores for younger-than-keyage boys and girls were aggregated over 7.00 to 7.99 years and that *delta z-scores* for older-than-keyage boys and girls were aggregated over 9.50 to 9.99 years. Points are binned *delta* child means. Coordination and speed times were converted from seconds to meters/seconds (i.e., pace scores; star-run test = 50.912 [m] / time [s]; 20-m linear sprint test = 20 [m] / time [s]). These transformations have the advantage that a large value is indicative of better physical fitness and that they remove skew in the distributions.

Physical fitness of older-than-keyage children (OTK)

Table 2 displays statistics for fixed effects of age (linear and quadratic) and sex as well as their interactions with the four test contrasts for LMM of OTK children.

Table 2 Fixed-effect estimates of linear mixed model for older-than-keyage (OTK) children

Source of variance	Fixed-effect estimates	Standard error	z-values
Main effects			
Grand mean (intercept)	0.348	0.068	5.08*
H1: coordination vs. endurance	0.230	0.091	2.51*
H2: speed vs. coordination	-0.025	0.083	-0.30
H3: powerLOW vs. speed	0.029	0.075	0.39
H4: powerUP vs. powerLOW	-0.022	0.094	-0.23
Age (linear)	-1.014	0.152	-6.68*
Age (quadratic)	0.357	0.079	4.50*
Sex	0.015	0.011	1.43
Age (linear) x Fitness component			
H1: coordination vs. endurance	-0.693	0.205	-3.38*
H2: speed vs. coordination26	0.180	0.187	0.96
H3: powerLOW vs. speed	-0.181	0.168	-1.08
H4: powerUP vs. powerLOW	0.089	0.211	0.42
Age (quadratic) x Fitness component			
H1: coordination vs. endurance	0.294	0.108	2.73*
H2: speed vs. coordination	-0.034	0.098	-0.34
H3: powerLOW vs. speed	0.059	0.088	0.67
H4: powerUP vs. powerLOW	-0.014	0.111	-0.13
Sex x Fitness component			
H1: coordination vs. endurance	0.052	0.015	3.41*
H2: speed vs. coordination	-0.001	0.014	-0.10
H3: powerLOW vs. speed	0.042	0.012	3.38*
H4: powerUP vs. powerLOW	-0.044	0.016	-2.80*

H1 to H4 = hypothesis 1 to 4, endurance = cardiorespiratory endurance (i.e., 6 min run test), coordination = star run test, speed = 20-m linear sprint test, powerLOW = lower limbs muscle power (i.e., standing long jump test), powerUP = upper limbs muscle power (i.e., ball push test), * = z-value > 2.0, linear mixed model random factors: assessment years (9), schools (513), children (26,540), observations = 128,198 (missing = 3.4%). For estimates of variance components and correlation parameters see Table 3.

The overall negative linear trend for age (z = -6.68) and positive quadratic trend of age (z = 4.50) were significant. The positive quadratic trend of age indicates that the difference between predicted and observed physical fitness becomes more negative initially, but plateaus with even a slight reduction of delta z-scores for the oldest children (see Figure 2).

Furthermore, the main effect of contrast H1 was significant (z = 2.51) indicating that the main effect was larger for coordination than for cardiorespiratory endurance. The LMM tested the interactions of linear and quadratic age with the four test contrasts, that is whether slopes in neighboring panels in Figure 2 (averaged across sex) were parallel. The slope can be equated with the developmental rate. Indeed, one of four interaction was significant (see second and third block of Table 2) the linear age developmental rate was larger for cardiorespiratory endurance than coordination (H1; z = -3.38) and the quadratic age developmental rate was larger for coordination than cardiorespiratory endurance (H1; z = 2.73).

Three of the test contrasts interacted with sex. First, the delta z-score was more negative for boys than girls for cardiorespiratory endurance and more negative for girls than boys for coordination (z = 3.41, see Table 1). The post-hoc LMM revealed significantly less severe delta z-scores for girls (-0.14) than boys (-0.18) for cardiorespiratory endurance (z= - 2.30). There was no significant sex difference for the delta z-score for coordination (z = 1.38). Second, the negative difference between boys and girls in the delta z-score was larger for powerLOW than speed (z = 3.38). The post-hoc LMM revealed a significant sex difference (favoring boys) only for powerLOW (z = 3.90; boys: -0.23, girls: -0.29; see Table 1). There was no significant sex difference for speed (z = 1.19). Third, the same powerLOW sex difference was the source of the significant interaction for the fourth contrast (z = -2,80). There was no significant sex difference for powerUP (z = 1.12).

Table 3 Variance components for older-than-keyage (OTK) children

Fitness component	GM	Endurance	Coordination	Speed	PowerLOW	PowerUP	Age
Assessment year	0.03						
School		0.08	0.07	0.07	0.06	0.07	0.09
Child (OTK)		0.92	0.94	0.94	0.98	0.88	

Endurance = cardiorespiratory endurance (i.e., 6 min run test), Coordination = star run test, Speed = 20-m linear sprint test, PowerLOW = lower limbs muscle power (i.e., standing long jump test), PowerUP = upper limbs muscle power (i.e., ball push test), OTK = older-than-keyage children. VC for Residual = 0.65.

Table 3 lists estimates of VCs for children and for school. The delta z-scores VCs were large for children (0.88 to 0.94) and small for schools (0.06 to 0.09).

Physical fitness of younger than keyage (YTK) children

Table 4 displays estimates and test statistics for fixed effects of the four test contrasts. Figure 2 displays the delta z-scores between observed and predicted physical fitness development for YTK boys and girls aggregated over 7 years and 0 months to 7 years and 11 months.

Table 4 Fixed-effect estimates of the linear mixed model for younger-than-keyage (YTK) children

Source of variance	Fixed-effect estimates	Standard error	z-values
Main effects			
Grand mean (intercept)	0.082	0.016	5.09*
H1: coordination vs. endurance	0.074	0.024	3.05*
H2: speed vs. coordination	-0.055	0.022	-2.49*
H3: powerLOW vs. speed	0.026	0.021	1.27
H4: powerUP vs. powerLOW	0.060	0.024	2.47*

H1 to H4 = hypothesis 1 to 4, endurance = cardiorespiratory endurance (i.e., 6 min run test), coordination = star run test, speed = 20-m linear sprint test, powerLOW = lower limbs muscle power (i.e., standing long jump test), powerUP = upper limbs muscle power (i.e., ball push test), * = z-value > 2.0, linear mixed model random factors: schools (437), children (2,586), observations = 12,590 (missing = 2.6%). For estimates of variance components and correlation parameters see Table 5.

The grand mean was significant (z = 5.09). Furthermore, three of the four main effects of contrasts were significant: the main effect was larger for coordination than cardiorespiratory endurance (H1; z = 3.05), larger for coordination than speed (H2; z = -2.49) and larger for powerUP than powerLOW (H4; z = 2.47), which can also be seen in Figure 2.

Table 5 Variance components for younger-than-keyage (YTK) children

Fitness component	Endurance	Coordination	Speed	PowerLOW	PowerUP
School	0.12	0.10	0.09	0.10	0.10
Child (YTK)	0.83	0.88	0.89	0.89	0.79

Endurance = cardiorespiratory endurance (i.e., 6 min run test), Coordination = star run test, Speed = 20-m linear sprint test, PowerLOW = lower limbs muscle power (i.e., standing long jump test), PowerUP = upper limbs muscle power (i.e., ball push test), YTK = younger-than-keyage (YTK) children. VC for Residual = 0.62.

Table 5 lists estimates of VCs between delta z-scores for children and for school. The delta z-scores VCs were large for children (0.83 to 0.89) and small for schools (0.09 to 0.12).

Discussion

The aim of this cross-sectional study was to examine physical fitness of YTK and OTK children versus keyage children in a representative sample of German primary school children. Our findings indicate that (i) OTK children showed poorer performance compared to keyage children, especially for coordination, (ii) OTK girls outperformed OTK boys, and (iii) YTK children showed better results than keyage children, especially for coordination.

Several studies confirmed a linear increase in physical fitness performance with chronological age[9–11,15]. For instance, in a study with 424,328 Greek children and adolescents aged 6 to 18 years, Tambalis et al.[15] reported a linear increase in physical fitness performance with age for cardiorespiratory endurance (i.e., 20-m shuttle run test), lower limbs muscle power (i.e., standing long jump test), flexibility (i.e., sit-and-reach test), muscular strength (i.e., sit-ups test), and agility (i.e., 10 × 5 m shuttle run test). The development of physical fitness of keyage children (see predicted gray lines in Figure 1) is in accordance with the above reported results. For keyage children, physical fitness performance increased linearly with age. However, the development of physical fitness for OTK children is different. Poor performance was found in OTK children aged 9 years and 0 months to 9 years and 11 months compared with age-matched keyage children for all components of physical fitness, especially for coordination. This could be due to the fact that third graders aged 9 years and 0 months to 9 years and 11 months (i.e., OTK children) are not representative for the “average” age-matched keyage child which is why we observed a deviation from the typically reported fitness development with age in this cohort[9–11,15]. We do not know the exact circumstances which lead to the delayed enrollment into first grade or to the repetition of a school year. According to our results, we can only speculate that maybe a delay in cognitive development might be the reason why children are late enrolled into first grade or must repeat a school class. These results are in line with a study of Urschitz et al.[23] who examined differences in academic performance. These authors observed that poor academic performance significantly increased with age for mathematics, science, reading, spelling, and handwriting in a sample of 1,144 German third graders. Of note, children who repeated a school class were more prone to poor academic performance. These results were confirmed by other studies for academic performance[21,22]. Interestingly, in our study OTK girls showed better performance compared to OTK boys which is in accordance with Urschitz et al.[23]. These authors reported that except for mathematics, boys showed a larger prevalence for poor academic performance compared with girls[23]. As girls mature approximately two years earlier than boys, the better performance of girls compared to boys might be influenced by biological maturation. Girls enter the adolescent growth spurt at approximately ten years of age and peak height velocity at 12 years, whereas boys enter the growth spurt on average at age 12 and peak height velocity at 14[36].

In contrast, YTK children outperformed keyage children especially in tests requiring motor coordination. Again, we do not know the exact circumstances which resulted in early enrollment into first grade or reasons for skipping a school year. According to our results, we speculate that accelerated cognitive development could be a reason why early enrolled children skip a school year. This is supported by the fact that in this study, YTK children showed the best performance in the coordination test which has an inherent large cognitive demand. Moreover, findings from Martin[22] point in a similar direction by showing that in a cohort of 3,684 Australian high school students, YTK children outperformed keyage children in academic performance.

Our study is not without limitations. First, anthropometric factors such as body mass, body height, and sitting height were not assessed in this study so that associations between anthropometric factors, biological maturation, and physical fitness could not be calculated. These factors would have provided additional insight as there is strong evidence that children's physical fitness is associated with anthropometric characteristics[37–39] and biological maturation[36]. One explanation of the deviation of YKT and OKT children might be a difference between chronological and biological age. It appears plausible to argue that YKT children may be more mature and that OKT children are biologically somewhat younger than indicated by their chronological age. Thus, in a hypothetical plot of performance over biological age, the linear trend may well hold for all children. Second, we predicted the performance of the YTK and OTK children based on a linear extrapolation recently reported by Fühner et al.[20]. However, we do not know if this linear extrapolation exactly fits to the data of keyage children aged 7 years and 0 months to 7 years and 11 months / 9 years and 0 months to 9 years and 11 months as we do not have such longitudinal data. Third, we cannot parse out the exact number of OTK children that were late enrolled or repeated a school class.

To sum up, this study is the first study that examined differences in physical fitness development of YTK and OTK children compared to keyage children. Our study findings complement results reported in the literature on the development of academic performance in youth[21–23]. Politicians and decision makers, schools, (physical education) teachers, and parents should be aware that OTK versus keyage children showed poorer physical fitness performance. This is a novel and somehow unexpected result. Therefore, OTK children should be specifically promoted through additional health and fitness programs to compensate their deficits in physical fitness to enable a holistic development. Furthermore, the assessment of physical fitness should be performed regularly to tailor the contents of physical education classes based on the results of physical fitness assessments (e.g., data driven physical education classes). More specifically, the physical fitness status of OTK children should be monitored regularly over time to evaluate whether e.g., additional health and fitness programs already helped to compensate the observed deficits in physical fitness.

Given that reference values for the grading of physical fitness is only available for keyage children, raw data from this study can be used to calculate age-, sex-, and timing of school enrollment-specific percentile values. The respective data should be useful for (physical education) teachers or researchers to individually evaluate and grade children´s physical fitness development.

The EMOTIKON test battery is easy-to-administer, cost effective, and it requires only minimal equipment that is usually available in gyms (e.g., stopwatch, measuring tape, medicine ball, pylons). Therefore, physical education teachers, coaches, or researchers can use the EMOTIKON test battery to evaluate children’s physical fitness and use the results to promote health- and skill-related physical fitness during physical education.

Data availability

The datasets generated and analyzed during the current study as well as Julia and R scripts are available in the Open Science Framework (OSF) repository: https://osf.io/dmu68/?view_only=240bdab8f1be4d8384acf9356ee50f8b

References

1. Ortega, F. B., Ruiz, J. R., Castillo, M. J. & Sjöström, M. Physical fitness in childhood and adolescence: A powerful marker of health. *Int. J. Obes*. **32**, 1–11 (2008).

2. Caspersen, C., Powell, K. & Christenson, G. Physical Activity, Exercise, and Physical Fitness: Definitions and Distinctions for Health-Related Research CARL. *Notes Queries* **100**(2), 125–131 (1985).

3. Hussey, J., Bell, C., Bennett, K., O'Dwyer, J. & Gormley, J. Relationship between the intensity of physical activity, inactivity, cardiorespiratory fitness and body composition in 7-10-year-old Dublin children. *Br. J. Sports Med*. **41**, 311–316 (2007).

4. Mintjens, S. et al. Cardiorespiratory Fitness in Childhood and Adolescence Affects Future Cardiovascular Risk Factors: A Systematic Review of Longitudinal Studies. *Sport. Med.* **48**, 2577–2605 (2018).

5. García-Hermoso, A., Ramírez-Campillo, R. & Izquierdo, M. Is Muscular Fitness Associated with Future Health Benefits in Children and Adolescents? A Systematic Review and Meta-Analysis of Longitudinal Studies. *Sport. Med*. **49**, 1079–1094 (2019).

6. Fühner, T., Kliegl, R., Arntz, F., Kriemler, S. & Granacher, U. An Update on Secular Trends in Physical Fitness of Children and Adolescents from 1972 to 2015: A Systematic Review. *Sport. Med*. **51**, 303–320 (2021).

7. Tomkinson, G. R., Lang, J. J. & Tremblay, M. S. Temporal trends in the cardiorespiratory fitness of children and adolescents representing 19 high-income and upper middle-income countries between 1981 and 2014. *Br. J. Sports Med*. **53**, 478–486 (2019).

8. Niessner, C. et al. Representative Percentile Curves of Physical Fitness From Early Childhood to Early Adulthood: The MoMo Study. *Front. Public Heal*. **8**, 1–9 (2020).

9. Tomkinson, G. R. et al. European normative values for physical fitness in children and adolescents aged 9-17 years: Results from 2 779 165 Eurofit performances representing 30 countries. *Br. J. Sports Med*. **52**, 1445–1456 (2018).

10. De Miguel-Etayo, P. et al. Physical fitness reference standards in European children: The IDEFICS study. *Int. J. Obes*. **38**, S57–S66 (2014).

11. Santos, R. et al. Physical fitness percentiles for Portuguese children and adolescents aged 10-18 years. *J. Sports Sci*. **32**, 1510–1518 (2014).

12. Ortega, F. B. et al. Physical fitness levels among European adolescents: The HELENA study. *Br. J. Sports Med*. **45**, 20–29 (2011).

13. Catley, M. J. & Tomkinson, G. R. Normative health-related fitness values for children: Analysis of 85347 test results on 9-17-year-old Australians since 1985. *Br. J. Sports Med.* **47**, 98–108 (2013).

14. Roriz de Oliveira, M. S., Seabra, A., Freitas, D., Eisenmann, J. C. & Maia, J. Physical fitness percentile charts for children aged 6-10 from Portugal. *J. Sports Med. Phys. Fitness* **54**, 780–792 (2014).

15. Tambalis, K. D. et al. 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. *Eur. J. Sport Sci*. **16**, 736–746 (2016).

16. Castro-Piñero, J. et al. Percentile values for running sprint field tests in children ages 6-17 years: Influence of weight status. *Res. Q. Exerc. Sport* **81**, 143–151 (2010).

17. Dobosz, J., Mayorga-Vega, D. & Viciana, J. Percentile values of physical fitness levels among polish children aged 7 to 19 years – a population-based study. *Cent. Eur. J. Public Health* **23**, 340–351 (2015).

18. Cobley, S., Baker, J., Wattie, N. & McKenna, J. Annual age-grouping and athlete development: A meta-analytical review of relative age effects in sport. *Sport. Med*. **39**, 235–256 (2009).

19. Musch, J. & Grondin, S. Unequal Competition as an Impediment to Personal Development: A Review of the Relative Age Effect in Sport. *Dev. Rev*. **21**, 147–167 (2001).

20. Fühner, T., Granacher, U., Golle, K. & Kliegl, R. Age and sex effects in physical fitness components of 108,295 third graders including 515 primary schools and 9 cohorts. *Sci. Rep*. **11**, (2021).

21. Jaekel, J., Strauss, V. Y. C., Johnson, S., Gilmore, C. & Wolke, D. Delayed school entry and academic performance: A natural experiment. *Dev. Med. Child Neurol.* **57**, 652–659 (2015).

22. Martin, A. J. Age Appropriateness and Motivation, Engagement, and Performance in High School: Effects of Age Within Cohort, Grade Retention, and Delayed School Entry. *J. Educ. Psychol*. **101**, 101–114 (2009).

23. Urschitz, M. S. et al. Snoring, intermittent hypoxia and academic performance in primary school children. *Am. J. Respir. Crit. Care Med*. **168**, 464–468 (2003).
24. Ministerium für Bildung Jugend und Sport. Gesetz über die Schulen im Land Brandenburg. https://bravors.brandenburg.de/gesetze/bbgschulg (2018).
25. World Medical Association. Declaration of Helsinki - Ethical Principles for Medical Research involving human Subjects. https://www.wma.net/policies-post/wma-declaration-of-helsinki-ethical-principles-for-medical-research-involving-human-subjects/ (2013).
26. Nuzzo, J. L. The Case for Retiring Flexibility as a Major Component of Physical Fitness. *Sport. Med*. **50**, 853–870 (2020).
27. Bös, K. *Deutscher Motorik-Test 6-18*. (Czwalina, 2009).
28. Schulz, S. *The reliability of the Star coordination run and the 1-kg medicine ball push - physical fitness tests used in the EMOTIKON-study*. (University of Potsdam, 2013).
29. Fernandez-Santos, J. R., Ruiz, J. R., Cohen, D. D., Gonzalez-Montesinos, J. L. & Castro-Pinero, J. Reliability and validity of tests to assess lower-body muscular power in children. *J. Strength Cond. Res*. **29**, 2277–2285 (2015).
30. Team, R. C. R a language and environment for statistical computing. https://www.r-project.org/ (2020).
31. Wickham, H. et al. Welcome to the Tidyverse. *J. Open Source Softw*. **4**, 1686 (2019).
32. Bates, D. et al. JuliaStats/MixedModels.jl:v4.6.0. vol. 1 20071787–20071787 https://doi.org/10.5281/zenodo.5825693 (2022).
33. Bezanson, J., Edelman, A., Karpinski, S. & Shah, V. B. Julia: A fresh approach to numerical computing. *SIAM Rev*. **59**, 65–98 (2017).
34. Box, G. E. P. & Cox, D. R. An analysis of transformations. *J. R. Stat. Soc. Ser. B* **26**, 211–243 (1964).
35. Bates, D., Kliegl, R., Vasishth, S. & Baayen, H. Parsimonious Mixed Models. ArXiv:1506.04967 http://arxiv.org/abs/1506.04967 (2015).
36. Malina, R. M., Bouchard, C. & Bar-Or, O. *Growth, Maturation, and Physical Activity*. (Human Kinetics, 2004).

37. Ceschia, A. et al. Deleterious effects of obesity on physical fitness in pre-pubertal children. *Eur. J. Sport Sci.* **16**, 271–278 (2016).

38. Sacchetti, R. et al. Physical fitness of primary school children in relation to overweight prevalence and physical activity habits. *J. Sports Sci*. **30**, 633–640 (2012).

39. Deforche, B. et al. Physical fitness and physical activity in obese and nonobese Flemish youth. *Obes. Res*. **11**, 434–441 (2003).

Funding

Funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – Projektnummer 491466077. Open Access funding enabled and organized by Projekt DEAL. The study was commissioned and supported by the Ministry of Education, Youth, and Sport of the Federal State Brandenburg. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript. Reinhold Kliegl was supported by the Center for Interdisciplinary Research, Bielefeld (ZiF)/Cooperation Group "Statistical models for psychological and linguistic data".

Acknowledgement

The authors thank Paula Teich for helpful comments.

Author contributions

TF, RK, and UG: made substantial contributions to conception and design; KG and TF: contributed to data collection; TF and RK: carried out data analysis; TF, RK, and UG: interpreted the data; TF: wrote the first draft of the manuscript and all authors were involved in revising it critically for important intellectual content; all authors provide final approval of the version to be published and agreed to be accountable for all aspects of the work and agreed with the order of presentation of the authors.

Competing Interests Statement

The authors declare no competing interests.

Curriculum Vitae

Page 157 (CV) contains personal data. It is therefore not part of the online and print publication.

Complete publication record

Original and review articles

Fühner, T., Granacher, U., Golle, K. & Kliegl, R. (2022). Effect of timing of school enrollment on physical fitness in third graders. *Scientific Reports*, *11*, 7801.

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*, 17566.

Haile, S.R., **Fühner, T.**, Granacher, U., Stocker, J., Radtke, T. & Kriemler, S. (2021). Reference values and validation of the 1-minute sit-to-stand test in healthy 5-16-year-old youth: a cross-sectional study. *BMJ Open, 11*, e049143.

Fühner, T., Kliegl, R., Arntz, F., Kriemler, S. & Granacher, U. (2021). An update on secular trends in physical fitness of children and adolescents from 1972 to 2015: A systematic review. *Sports Medicine, 51*, 303-320.

Fühner, T., Gebel, A. & Granacher, U. (2020). Entwicklung und Förderung von Kraft und Gleichgewicht von Kindern und Jugendlichen unter Berücksichtigung des Sportunterrichts: ein narrativer Überblicksbeitrag. *Leipziger Sportwissenschaftliche Beiträge, 61* (1), 29-50.

Ferrauti, A., Fett, J., Vuong, J.L., **Fühner, T.** & Ulbricht, A. (2019). Zehn Jahre DTB-Konditionstest. Ein Rückblick. *Tennissport Fachzeitschrift für Training und Wettkampf, 2*, 12-19.

Book chapters

Fühner, T., Kriemler, S., Woll, A. & Granacher, U. (2021). Körperliche Leistungsfähigkeit im Kontext von Wachstum und Reifung. In I. Menrath, C. Graf, U. Granacher & S. Kriemler (Hrsg.), *Pädiatrische Sportmedizin*. *Kompendium für Kinder- und Jugendärzte, Hausärzte und Sportärzte* (pp. 23-33). Berlin: Springer.

Conference papers

Fühner, T., Granacher, U., Golle, K., Teich, P. & Kliegl, R. (2022). Alters- und Geschlechtseffekte in der körperlichen Fitness. In M. Wegner & J. Jürgensen (Hrsg.), *Sport, Mehr & Meer – Sportwissenschaft in gesellschaftlicher Verantwortung, 25. dvs-Hochschultag an der Christian-Albrechts-Universität zu Kiel* (Schriften der Deutschen Vereinigung für Sportwissenschaft, 298). Hamburg: Czwalina.

Kliegl, R., **Fühner, T.** & Teich, P. (2022). Körperliche Fitness Brandenburger Drittklässler vor und seit Covid-19. In M. Wegner & J. Jürgensen (Hrsg.), *Sport, Mehr & Meer – Sportwissenschaft in gesellschaftlicher Verantwortung, 25. dvs-Hochschultag an der Christian-Albrechts-Universität zu Kiel* (Schriften der Deutschen Vereinigung für Sportwissenschaft, 298). Hamburg: Czwalina.

Fühner, T., Fett, J., Vuong, J.L., Ulbricht, A. & Ferrauti, A. (2019). Bedeutung der konditionellen Leistungsfaktoren für das Erreichen der nationalen und internationalen Spitzenklasse im Tennis. In A. Arampatzis, S. Braun, K. Schmitt & B. Wolfarth (Hrsg.), *Sport im öffentlichen Raum, 24. dvs-Hochschultag an der Humboldt-Universität zu Berlin* (Schriften der Deutschen Vereinigung für Sportwissenschaft, 282). Hamburg: Czwalina