# A study to examine the validity and reliability of a selected group of physical fitness tests within a primary school setting

*Author:*

Mr Robin Buck

*Supervisor:*

Professor Mike Lambert

*Submitted in fulfilment of the requirements for the*

*Degree of Master of Exercise Science*

Division of Physiological Sciences

Department of Human Biology

Faculty of Health Sciences

University of Cape Town

February 2025

**DECLARATION**

I, **Rob Buck**, hereby declare that the work on which this dissertation is based is my original work, except where acknowledgements indicate otherwise. I confirm that neither the whole work nor any part of it has been, is being, or is to be submitted for another degree in this or any other university.

I empower the university to reproduce for the purpose of research either the whole or any portion of the contents in any manner whatsoever.

Signed: Signed by candidate

# Table of Contents

## ACKNOWLEDGEMENTS

I wish to express my heartfelt gratitude to everyone who has contributed to the completion of this thesis. The journey to completion took a lot longer than I had originally anticipated. I would especially like to thank my thesis supervisor, Professor Mike Lambert, for his insightful guidance, unwavering support and incredible patience throughout the process from project conception, to publishing a paper, an "almost" upgrade to PhD, and completion of my dissertation.

My wife, Angela, and children, Liam and Kaity, were an immense support throughout the process. Choosing to complete an MSc while working full-time, has meant that all the time devoted to this project has come from "family time". I am incredibly grateful for their sacrifices and willingness to give me the time I needed and for their unconditional love, support, encouragement and steadfast belief in my abilities. I would also like to thank my parents for their ongoing support, intentional encouragement and belief in my capabilities.

I would like to offer a sincere appreciation to all the participants in the study. Without their willingness to run, jump and be measured, none of this would have been possible.

Finally, thank you God for your incredible love and grace, and for placing within me a curiosity to ask questions and a desire to look for answers.

## ABSTRACT

**Purpose:** This study represents an effort to provide an evidence-based approach to selecting field-based fitness measures for children and youth for inclusion in a test battery to be used within the school setting. The approach used by the author considers the following criteria to determine which youth fitness test items would be most appropriate for inclusion: (1) the strength of the fitness tests' association with health markers in youth, (2) the fitness tests integrity (validity and reliability), (3) the relative feasibility of the test administration in the field and (4) the measurement error of the fitness test. This study aims to identify valid and reliable physical fitness tests with low measurement errors for use in South African schools

**Method:** A literature review was performed to determine the predictive and criterion validity of physical fitness tests in children. Tests with strong predictive evidence and good criterion validity were selected for further reliability testing. A group of 290 male and female children aged 6 to 13 years performed a physical fitness testing protocol three times within a 7-day period. The typical error of measurement (TE), coefficient of variation ($CV_{TE}$) and smallest worthwhile change (SWC) were calculated for each sex and age group (grade). Changes in measurement were categorised as *small*, *medium* or *large* based on effect sizes.

**Results:** The 20m shuttle run (SRT), hand grip strength (HGS), standing long jump (SLJ), body mass index (BMI), and waist circumference (WC) tests demonstrate strong predictive evidence and good criterion validity. Typical error for the 20mSRT was 0.9 stages. *Small* changes in 20mSRT (0.4 stages) are challenging for the test to detect. *Medium* (1.0 stages) and *large* (1.9 stages) changes in the 20mSRT are more interpretable as they exceed the "noise" (TE) of the measurement. *Small (0.6* kg/m$^2$*), medium* (1.6 kg/m$^2$) and *large* (2.5 kg/m$^2$) changes in BMI are all larger than the TE (0.2 kg/m$^2$). The BMI test has a 'good' ability to detect *small, medium* and *large* changes. TE for HGS was 1.3 kg. *Small* changes in HGS (1.3kg) are difficult for the HGS test to detect. *Medium* (3.3kg) and *large* (5.3kg) changes in HGS are more interpretable as they exceed the "noise" (TE) of the measurement. The WC test has a 'good' ability to detect *small (1.5* cm*),* *medium* (3.8 cm) and *large* (6.1 cm) changes as they are larger than the TE (1.5 cm). TE for SLJ was 8.8 cm.

*Small* changes in SLJ (5.4 cm) are difficult for the SLJ test to detect. *Medium* (13.5 cm) and *large* (21.5 cm) changes in SLJ are more interpretable as they exceed the “noise” (TE) of the measurement.

**Conclusion:** The field-based physical fitness tests that demonstrated strong predictive validity and good criterion validity from the literature and low measurement error from the reliability data are HGS, SLJ, BMI, WC and 20mSRT. These fitness tests could be included in a battery of tests for implementation within the school setting. Researchers and practitioners can use the SWC and TE as guidelines to set targets when determining the extent to which performance changes in these fitness tests are practically significant.

# CHAPTER 1: OVERVIEW

## Thesis Publications

Buck R, Lambert MI. Getting a Grip on Strength Measurement in Children (6-13 Y): Impact of Typical Error of Measurement. Pediatr Exerc Sci. 2022 Feb 2;34(3):141-147. doi: 10.1123/pes.2021-0069. PMID: 35108675.

## List of Tables and Abbreviations

## List of Tables

## List of Abbreviations

ANOVA – analysis of variance

BMI – body mass index

CI – confidence interval

CM – centimetre

CV – coefficient of variation

CVD – cardiovascular disease

F – female

HGS – hand grip strength

KG – kilogram

KM – kilometre

M – male

N – sample size

P – probability

RM – repetition maximum

SD – standard deviation

SLJ – standing long jump

SRT – shuttle run test

SWC – smallest worthwhile change

TE – typical error of measurement

$VO_2max$ – maximal oxygen uptake

WC – waist circumference

## General Introduction (including aims and objectives of the study)

Physical fitness reflects an individual's ability to perform specific exercises or functions and is related to present and future health outcomes (Institute of Medicine, 2012). It is typically defined as a set of attributes that are either health or skill/performance-related. Health-related physical fitness describes the ability of a person to perform daily activities with vigour (Ortega et al., 2008; Ruiz et al., 2009). It is characterised by attributes associated with a low risk of developing chronic diseases and premature death (Ortega et al., 2008; Ruiz et al., 2009). Chronic and cardiovascular disease events most frequently occur during or after the fifth decade of life. However, evidence indicates that the precursors of cardiovascular disease originate in childhood and adolescence (Ortega et al., 2008; Ruiz et al., 2009). Many researchers have suggested that chronic disease and adult obesity have their origins in childhood inactivity (Batty and Leon, 2002; Chatrath, Shenoy, Serratto, and Thoele, 2002; Tomkinson, Leger, Olds, and Cazorla, 2003).

Physical fitness in youth has thus been recognised as a significant health-related indicator (Ortega, et al., 2008; Ruiz, et al., 2009; Minatto et al., 2016; Mintjens et al., 2018; García-Hermoso et al., 2020), and the assessment of health-related fitness in childhood and adolescence is of public health and clinical interest. Truter, et al. (2010), showed an inverse relationship between physical activity levels and obesity and overweight in South African children. Furthermore, a study by Ruiz, et al. (2010) demonstrated that higher levels of cardiorespiratory fitness, muscular strength improvements and a healthier body composition during childhood and adolescence are associated with a healthier cardiovascular profile and body composition later in life.

The education system in South Africa, particularly at the primary school level, reflects the country's complex history and ongoing social and economic challenges. While the system is governed nationally by the

Department of Basic Education (DBE), the quality and delivery of education vary widely between provinces, urban and rural areas, and schools with different socio-economic profiles.

While the education system is built on strong policy frameworks, there is an implementation gap that is affected by the inequality, poverty, and historical imbalances and this poses major challenges. These issues not only affect academic achievement but also have a direct impact on physical education and the feasibility of implementation of sport programmes. The stark contrast between well-resourced schools and under-resourced schools in rural areas or townships means that many children are not exposed to sufficient levels of physical activity.

This reinforces the need to identify approaches to both promote physical activity and ensure ongoing monitoring and surveillance of South African children. Identifying children at risk of poor health outcomes is crucial. Once high-risk children are identified, interventions can be implemented and monitored with follow-up physical fitness testing. This enables those children who present as a risk to have the opportunity of a healthier future. This makes the process of selecting physical fitness tests that are most appropriate for implementation in the school system that much more important.

In the 2018 HAKSA Report Card, it was estimated that about half of South African children and adolescents met the global recommendations for overall physical activity (HAKSA, 2018). However, at the same time, nearly one-third of children were not participating in regular, school-based physical education (Silva et al, 2018). It was further reported that there is a gap between the existing policies and the actual implementation of these policies (HAKSA, 2018; Nyawose et al, 2024). The 2022 HAKSA report card showed an improvement in physical activity indicators including overall physical activity, physical fitness, and government policy and programs (HAKSA, 2022). However, body composition proxies and most of the nutrition indicators remained unchanged (HAKSA, 2022). The indicators that regressed from 2018 to 2022 included community and environmental influences, as well as participation in organized sport (Nyawose et al, 2024). Despite the apparent improvement in overall physical activity levels in children and adolescents, there is a lack of tangible evidence of actual implementation of policies and programs (Nyawose et al, 2024). Overall, there is a need to identify approaches to promote physical activity in South African children and adolescents and ongoing monitoring and surveillance is crucial (Nyawose et al, 2024). Therefore, the physical fitness levels of South African children need to be assessed and monitored to identify children at risk of poor health outcomes. Once high-risk children are identified, interventions can be implemented and monitored with follow-up physical fitness testing. This enables those children who present as a risk to have the opportunity of a healthier future.

Laboratory methods have been used to objectively and accurately measure physical fitness (Bassett, 2001; Milliken et al., 2008; Castro-Piñero, 2009). However, these tests are not yet feasible at a population level due to their high cost, time constraints and the need for sophisticated instruments and qualified technicians (Ruiz et al., 2011). On the other hand, field-based fitness tests offer a more feasible and practical alternative. They are easy to administer, involve minimal equipment, are low-cost, and can evaluate large numbers of participants in a relatively short period (Paineau et al., 2008; Ruiz et al., 2008; España-Romero et al., 2010). Since most children and adolescents attend school, it is an appropriate setting to collect physical fitness data. These data can be used to identify children with low physical fitness and poor health outcomes, and interventions can be implemented to encourage positive fitness-enhancing behaviours. Assessing physical fitness levels through standardised field tests at schools where most children and adolescents can be reached provides an opportunity to monitor health on a population level.

The accuracy of interpreting measurements from physical fitness tests depends on knowing their reliability, sensitivity, and the magnitude of change representing practical significance. Lessons about interpreting measurements can be learnt from other disciplines. For example, in high-level sports, the interpretation of measurements to affect decision-making about athlete management is refined (Gabbett et al., 2017). Athletes are monitored in many sports to determine whether they adapt to the training load. The support staff of these athletes need to know whether an improvement in a measurement following any intervention is a real change and of practical significance or merely due to measurement error (Roe et al., 2016). The measurement error can be influenced by familiarity, warm-up, motivation, equipment, and environmental conditions. Controlling these factors affects how the measurement is interpreted (Hopkins, 2004).

Suppose that practically significant changes in the measurement (i.e., the signal) are small. In that case, the typical error of the measurement (i.e., the noise) must be low to ensure the signal can be detected above the noise. The noise/signal ratio of the measurement can be represented as the ratio of the typical error of measurement and the smallest worthwhile change (Gabbett et al., 2017). If the typical error exceeds the smallest worthwhile change, the measurement will only detect large changes which are meaningful. Conversely, small changes in the measurement will be detected if the typical error is lower than the smallest worthwhile change. This approach is well established in high-performance sports (Roe et al., 2016) and should be adopted when monitoring and analysing physical fitness test measurements of children and adolescents. It improves the measurement's interpretability and adds value when assessing differences between groups or deciding whether intervention programs have been effective.

Although a few studies have examined the validity and reliability of physical fitness tests, not many have examined these issues within the South African context. The author is also unaware of any research that has examined the typical error of measurement and the smallest worthwhile change of physical fitness tests among South African children. Researchers and practitioners need to have insight into the practical relevance of the observed changes in fitness measurements. Knowing and understanding the typical error ("noise") of a measurement (TE) and the smallest worthwhile change (SWC), allows the practitioner an extra level of analysis when reviewing changes in physical fitness test measurements (Buck and Lambert, 2022; Swinton et al., 2018).

The research that follows is an attempt to provide the practitioner with a battery of appropriate physical fitness tests and a means to help identify what performance change is considered of practical significance (Moir et al., 2004; Spencer et al., 2006) by knowing the SWC and TE for each of these tests. Researchers and practitioners could utilise the SWC and TE as guidelines for setting targets when evaluating the degree of certainty concerning the practical significance of performance changes in a physical fitness characteristic (Buck and Lambert, 2022; Roe et al., 2016; Swinton et al., 2018). Numerous efforts have been made to identify suitable fitness tests and to create standardised batteries for such tests for children and youth. Various historical circumstances and interpretations of the science have resulted in the use of many different fitness test batteries within the school setting (Ruiz et al., 2010; Marques et al., 2021). This study represents an effort to provide an evidence-based approach to selecting field-based fitness measures for children and youth for inclusion in a test battery to be used within the school setting. The approach used by the author considers the following criteria to determine which youth fitness test items would be most appropriate for inclusion:

(1) the strength of the fitness tests' association with health markers in youth,

(2) the fitness tests integrity (validity and reliability),

(3) the relative feasibility of the test administration in the field and

(4) the measurement error of the fitness test

Therefore, this study aims to determine which physical fitness field tests are most appropriate to utilise within a South African primary school setting based on their reported validity, reliability data and low measurement error.

## Purpose Of The Study

**1. Aim:**

To measure the reliability of physical fitness tests for children to determine which of these tests are appropriate for inclusion in a programme for use in a primary school setting.

**2. Research Question:**

What is the measurement error of children's physical fitness tests? Which tests with good validity and low measurement error should be incorporated into a programme at primary schools?

**3. Objectives:**

**3.1.** Perform a literature review that includes:

- the predictive validity of the physical fitness tests performed within a school setting
- the criterion validity of the field-based tests performed within a school setting

**3.2.** Examine the reliability of selected field-based tests

- Calculate the typical error of measurement and smallest worthwhile change of each measurement for each age group and sex
- Calculate the average typical error of measurement and smallest worthwhile change for the primary school population
- Determine which field tests are the most appropriate to use within a school setting based on the reliability data and reported validity

# CHAPTER 2: REVIEW OF LITERATURE

In recent decades, the fitness of children and adolescents has received considerable attention (Castro-Piñero et al., 2009 and Marques et al., 2021). Consequently, numerous field-based fitness test batteries have been developed to assess physical fitness in this population. Table 1 summarises 19 different physical fitness test batteries from European, American, Asian, and Oceanian countries used to assess health-related fitness in children aged 6 to 13.

***Table 1:*** *Field-based physical fitness test batteries that include children within the age bracket of 6 – 13 years*

| Age (years) | Fitness Test Battery | Setting in which utilised | Country/Region |
|---|---|---|---|
| 6-18 | ALPHA | School, sports and health | Europe |
| 6–18 | EUROFIT | School and health | Europe |
| 5–17 | FITNESSGRAM | School, sports and health | United States of America |
| 6–17 | PCPF (President's Challenge: Physical Fitness) | School and health | United States of America |
| 6–17 | AAUTB (Amateur Athletic Union Test Battery) | School and sports | United States of America |
| 6–17 | YMCAYFT (YMCA Youth Fitness Test) | School and health | United States of America |
| 5–17 | NYPFP (National Youth Physical Program) | School, health and military | United States of America |
| 5–18 | HRFT (Health-Related Fitness Test) | School, health | United States of America |
| 5–18 | Physical Best | School, health | United States of America |
| 6–12 | NZFT (New Zealand Fitness Test) | School and health | New Zealand |
| 5-18 | AAHPER | School and health | United States of America |
| 6-18 | Bouge | School and health | France |
| 7-17 | CAHPER-FPT | School | Canada |
| 5-12 | Physical fitness test (PFTB) | School and health | Norway |
| 7-18 | INDARES | School and health | Czech Republic |
| 6-17 | PFAAT | School, sports and health | Japan |
| 6-17 | PROESP | School, sports and health | Brazil |

| Age (years) | Fitness Test Battery | Setting in which utilised | Country/Region |
|---|---|---|---|
| 6-19 | SLOfit | School and health | Slovenia |
| 6-60 | UNIFITTEST | School, sports and health | Czech Republic |

*Table adapted/compiled from Ruiz et al., 2010 and Marques et al., 2021*

There are 60 physical fitness tests within the 19 physical fitness testing batteries compiled from systematic reviews performed by Castro-Piñero et al., 2009 and Marques et al., 2021 and summarised in Table 1. These field-based tests aim to assess the various components of health-related fitness and skill/performance-related fitness, namely cardiorespiratory fitness, musculoskeletal fitness, body composition, and motor fitness. Understanding which physical fitness components in young people relate to future disease/better health status will help clarify which physical fitness characteristics need to be assessed and monitored to ensure better health outcomes.

When conducting physical fitness testing in children, the concepts of validity and reliability are crucial to ensure that the results are meaningful and useful. Validity refers to the degree to which a test actually measures what it claims to measure. In the context of children's fitness testing, this could involve ensuring that a 20m shuttle run test truly assesses cardiovascular endurance or that a standing long jump test genuinely reflects muscular strength and power.

There are several types of validity: criterion validity compares the test against a recognized standard (for example, $VO_2$ max testing for aerobic capacity); and construct validity verifies that the test aligns with theoretical concepts, such as developmental expectations for physical capabilities in different age groups. Predictive validity describes how well a fitness test can anticipate future outcomes and provides early indicators that can be used for intervention, such as identifying children at risk for obesity, cardiovascular disease, or inactivity.

Reliability, on the other hand, relates to the consistency and stability of test results over time. A reliable fitness test will yield similar results under consistent conditions, whether administered on different days (test-retest reliability) or by different assessors (inter-rater reliability). This is especially important with children, as their physical and psychological states can fluctuate significantly from day to day. Factors such as motivation, attention span, and even time of day can influence performance, so protocols must be standardized and age-appropriate to improve reliability.

Testing children can presents unique challenges. Children develop at varying rates, and chronological age does not always align with physical or emotional maturity. A fitness test must therefore be adaptable and sensitive enough to accommodate these differences. Moreover, ethical considerations arise: children should not be pressured or placed under undue stress during testing, and the purpose of testing must be clearly communicated to avoid anxiety or feelings of inadequacy. Additionally, environmental factors—such as equipment suitability, surface safety, and weather conditions—can influence test validity and reliability in school or community settings. Ensuring physical fitness tests are valid and reliable in children requires careful selection and adaptation of testing protocols. Test designers and administrators must consider developmental appropriateness, standardization of procedures, and ethical conduct to ensure that the results are both scientifically sound and practically meaningful for guiding health and physical education interventions.

It is impractical to utilise many different fitness tests within the school setting due to time and resource limitations. Knowing which fitness tests to prioritise is of crucial importance. Ruiz et al., 2011 demonstrated that for a physical fitness test to be included in a school physical fitness testing battery, the fitness test should demonstrate:

1) strong predictive validity,
2) good criterion validity and
3) feasibility of implementation (time-efficient, cost-effective, easy to understand and administer)

There is a need to identify those fitness tests that are most appropriate for use within the school setting based on their validity, reliability, feasibility, scalability and low measurement error (Safrit, 1986; Docherty, 1996; España-Romero et al., 2010; Ruiz et al., 2011; Institute of Medicine, 2012; Buck and Lambert, 2022; Washif et al., 2024; Cruz-León et al., 2025). What follows is a review of the literature on the predictive validity, criterion validity and feasibility of physical fitness tests used with children and adolescents. Within the literature, physical fitness tests have been described in two main categories: health-related physical fitness and skill or performance-related fitness. Cardiorespiratory fitness, muscular fitness, and body composition are referred to as health-related fitness and are often associated with disease prevention and health promotion, while balance, coordination, speed, agility and power are described as performance-related fitness, reflecting the performance aspect of physical fitness (Powell et al., 1998; Howley et al., 2001; Haga et al., 2008).

The literature review aims to identify those physical fitness tests that are most appropriate to utilise within the school setting based on their predictive validity, criterion validity and feasibility. The identified tests will then undergo further investigation to determine their typical error of measurement (TE) and smallest worthwhile change (SWC). This will provide researchers and practitioners with a focused battery of relevant fitness tests for use within the school system. Knowing the TE and SWC of these tests will allow an extra level of analysis when determining the practical relevance of the observed changes over time (Buck and Lambert, 2022).

## 2.1 Cardiorespiratory Fitness

Several longitudinal studies and systematic reviews (Ortega et al., 2008; España-Romero et al., 2010; Minatto et al., 2016; Mintjens et al., 2018; García-Hermoso et al., 2020) in children and adolescents report the relationship between physical fitness and the risk of developing an unhealthy cardiovascular profile later in life. In a systematic review by Ruiz et al. (2009) on the predictive validity of health-related fitness in youth, strong evidence supported that higher levels of cardiorespiratory fitness in childhood and adolescence are associated with a healthier cardiovascular profile later in life. This finding was supported by a cohort study of one million men by Henriksson and Ruiz et al. (2020). Cardiorespiratory fitness in childhood and adolescence predicts cardiovascular risk factors later in life, such as abnormal blood lipids, high blood pressure and overall and central adiposity (Ruiz et al.,2009; Henriksson and Ruiz et al., 2020).

A systematic review by Castro-Piñero et al. (2009) reported on measurements of cardiorespiratory fitness and triglycerides, the ratio of total cholesterol and high-density lipoproteins and body fat in a population of girls and boys. The upper quartile of this population was defined as being at risk. Participants were defined as "a case" if they had two or more risk factors. Over eight years, they found that the odds ratios for having two or more risk factors between quartiles of cardiorespiratory fitness were 3.1, 3.8 and 4.9 for quartiles two to four, respectively. At the second examination, odds ratios were 0.7, 3.5 and 4.9, respectively. The probability for "a case" at the first examination to be "a case" at the second was 6.0. In summary, this review showed that lower cardiorespiratory fitness was associated with higher odds of having multiple risk factors for cardiovascular health over eight years, with the likelihood increasing significantly across fitness quartiles and a strong persistence of risk factors over time.

Johnson et al. (2000) found that cardiorespiratory fitness was negatively associated with increased adiposity. Children with a higher cardiorespiratory fitness at the start of the study had a lower rate of increase in adiposity over the five-year study duration and follow-up. Moderate evidence exists to show that higher levels of cardiorespiratory fitness in childhood and adolescence can reduce the risk of developing metabolic syndrome and arterial stiffness later in life. There is also moderate evidence that demonstrates increasing cardiorespiratory fitness is inversely associated with changes in blood lipids and lipoproteins. Hasselstrom et al. (2002), Twisk et al. (2002) and Janz et al. (2002) found that cardiorespiratory changes were negatively correlated with changes in total cholesterol, triglycerides and the ratio of total cholesterol and high-density lipoproteins ($p<0.05$). In a study by Byrd-Williams et al. (2008), higher cardiorespiratory fitness was associated with less subsequent gain in body fat in boys ($\beta = -0.001$); $p=0.030$), but not in girls ($\beta = 0.0005$; $p=0.370$). However, Koutedakis et al. (2005), found that cardiorespiratory fitness was inversely associated with changes in body fat (β= -0.09); $p<0.050$) in both boys and girls. There is inconclusive evidence indicating that changes in cardiorespiratory fitness are associated with changes in carotid intima-media thickness, carotid distension and compliance, as well as weight gain, diabetes and metabolic syndrome (Ruiz et al., 2009). Since there is strong predictive evidence that cardiorespiratory fitness is related to future disease and/or better health status, physical fitness tests within this category should be considered good options for inclusion in a battery of tests to utilise within the school setting.

Cardiorespiratory fitness is the most studied component of physical fitness among children and adolescents (Falk, 2018). It is typically defined as the maximal oxygen uptake ($VO_2$max) that an individual can reach. The criterion measure for cardiorespiratory fitness is widely considered to be the $VO_2$max attained during a graded maximal exercise test in a laboratory-based setting (Pescatello et al., 2014). A portable gas analyser can also be worn during a field-based graded maximal exercise test (Castagna et al., 2010; Silva et al., 2012). However, measuring $VO_2$max using sophisticated and expensive instrumentation, laboratory technicians, and lengthy testing sessions in various settings, such as schools, sports clubs, and large-scale research studies, is neither practical nor feasible from a time and cost perspective (Pescatello et al.., 2014). In these settings, a useful alternative to estimate $VO_2$max, is the performance score obtained during cardiorespiratory fitness field tests (Mayorga-Vega et al., 2015).

Many field-based fitness tests have been proposed to assess cardiorespiratory fitness. Table 2 shows ten of these tests from nineteen fitness test batteries, including children aged 6 - 13 years (Ruiz et al., 2010; Marques et al., 2021).

***Table 2:*** *Field-based fitness tests utilised to assess cardiorespiratory fitness*

| **Fitness Test** | **Fitness test battery** | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | ALPHA | EUROFIT | FITNESSGRAM | PCPF | AAUTB | YMCAYFT | NYPFP | HRFT | Physical Best | NZFT | AAHPER | Bouge | CAHPER-FPT | PFTB | INDARES | PFAAT | PROESP | SLOfit | UNIFITTEST |
| ***Aerobic capacity*** | | | | | | | | | | | | | | | | | | | |
| 20m shuttle run | X | X | X | X | | | | | | | | X | | | X | X | | | |
| Hoosier endurance shuttle run | | | | | X | | | | | | | | | | | | | | |
| 1.5-mile run/walk test | | | | | | | | | X | | | | X | | | | | | |
| 1-mile run/walk test | | | X | X | X | X | X | X | X | | | | X | | | | | | |
| ½-mile run/walk test | | | | X | X | | | | X | | X | X | X | | | | | X | |
| ¼-mile run/walk test | | | | X | X | | | | | | | | | | | | | | |
| 12-minute run/walk test (Cooper test) | | | | | | | | X | | X | | | | X* | | | | | X |
| 1000m run | | | | | | | | | | | | | X | | | | | | |
| 1500m run/walk test | | | | | | | | | | | | | | | X | | | | |
| 6-minute run/walk test | | | | | | | | | | | | | | | | | X | | |

** Reduced Cooper test*
*Table adapted/compiled from Ruiz et al., 2010; Marques et al., 2021*

When considering which cardiorespiratory fitness test to utilise, this choice should be based on a test's validity and feasibility (Mayorga-Vega, 2015). According to meta-analyses by Mayorga-Vega et al., in 2015 and 2016, the 20m shuttle run (r = 0.78, 0.72-0.85), 1.5-mile walk/run (r = 0.79, 0.73-0.85), and 12-minute walk/run (r = 0.78, 0.72-0.83) show the greatest criterion-related validity for estimating cardiorespiratory fitness. These field-based tests demonstrate correlation coefficients that indicate moderate to high criterion validity for estimating oxygen uptake. Walk/run tests of shorter distances show poorer results of criterion validity. In contrast, longer tests do not demonstrate significantly higher criterion validity and could, therefore, be an unnecessary use of extra time and effort. The 20m shuttle run test demonstrates a higher criterion validity for children when combined with other variables such as age, sex, body mass or body mass index. This is not the case with walk/run tests. From a criterion validity perspective, both the walk/run tests (1.5 mile, 12-minute) and the 20m shuttle run test present useful methods to estimate cardiorespiratory fitness. España-Romero (2010) and Domone (2016) suggest that the feasibility and scalability of fitness tests are worth considering when deciding which test/s to utilise. Fitness tests that are easy to understand and simple to perform by children are more likely to produce valid and reliable results (Castro-Pinero et al., 2009; Ruiz et al., 2011) and be implemented (España-Romero et al., 2010; Domone et al., 2016; Washif et al., 2024) by schools.

Regarding assessing cardiorespiratory fitness in children, walk/run tests may not be suitable as they do not consider a child's ability to maintain an appropriate pace. Participants may start too fast and are therefore unable to maintain the speed throughout the test, or they may start so slow that when they want to increase the speed, the test is already finished. To improve this issue, Castro-Pinero et al. (2009) utilised various versions of the walk/run tests by altering the distance variable. Many of these different distances demonstrated poor criterion-related validity.

Ortega et al. (2008), España-Romero et al. (2010), Ruiz et al. (2011) and Tomkinson et al. (2019) demonstrated through several studies that the 20m shuttle run test (20mSRT) has sufficient safety, feasibility, validity, and reliability to be included in a battery of field-based fitness tests for implementation in a school setting. Their findings show moderate to high validity (r = 0.78) and high to very high reliability, with intra-class correlation coefficients ranging from 0.78 to 0.93 in children aged 8 – 18 years. Test-retest differences in means are negligible in children and youth. Cruz-León et al. (2025) also found in their systematic review that the 20mSRT demonstrated strong evidence for feasibility and moderate evidence for safety. Following the above review of the literature, the 20mSRT will be considered the most appropriate test for cardiorespiratory fitness within the primary school setting and will, as part of this study, undergo further testing to establish the typical error of measurement and smallest worthwhile change for each age group and sex.

## 2.2 Musculoskeletal Fitness

Musculoskeletal fitness is an important marker of health throughout life (Garcia-Hermoso et al., 2019). Low musculoskeletal fitness is recognised as a strong marker of a poor metabolic profile during childhood and adolescence. It is associated with several non-communicable diseases (Ortega et al., 2008), and with mortality in adulthood (Garcia-Hermoso et al., 2018). Strong evidence indicates that muscular strength changes from childhood to adolescence are negatively associated with changes in overall adiposity (Ortega, 2008; Henriksson, 2020). Strong evidence was found for an inverse relationship between musculoskeletal fitness and changes in central adiposity, cardiovascular disease, and metabolic risk factors (García-Hermoso et al., 2019). In the systematic review and meta-analysis by Smith et al. (2014), the pooled effect size for the relationship between musculoskeletal fitness and adiposity was r = -0.25 (95% CI -0.41 to -0.08).

A positive association also exists between musculoskeletal fitness and bone health (Smith et al., 2014; García-Hermoso et al., 2019). Inconclusive evidence indicates that muscular strength or flexibility changes in childhood or adolescence are predictors of musculoskeletal pain later in life. Where evidence of an association was found, these associations were low to moderate (Smith, 2014; García-Hermoso, 2019). The meta-analysis by Garcia-Hermoso et al. in 2019 demonstrated the following results:
A significant ($p < 0.050$), moderate-large association between muscular fitness and body mass index (r = − 0.14; 95% confidence interval (CI) − 0.21 to − 0.07), skinfold thickness (r = − 0.32; 95% CI − 0.40 to − 0.23), insulin resistance (estimated) (r = − 0.10; 95% CI − 0.16 to − 0.05), triglycerides (r = − 0.22; 95% CI − 0.30 to − 0.13), cardiovascular disease risk score (r = − 0.29; 95% CI − 0.39 to − 0.18), and bone mineral density (r = 0.166; 95% CI 0.086 to 0.243).

Low musculoskeletal fitness has also been associated with prediabetes or type 2 diabetes. Fraser et al., 2020 demonstrated that greater grip strength across the life course could protect against prediabetes and type 2 diabetes. The authors found that a one standard deviation increase in cumulative grip strength was associated with 34% reduced odds of prediabetes or type 2 diabetes in mid-adulthood (odds ratio = 0.66, 95% CI: 0.40 - 0.98). Including physical fitness tests within the musculoskeletal fitness category seems reasonable based on the strong predictive evidence relating to future disease and/or better health status. Table 3 shows 28 different tests that have been utilised to assess different aspects of musculoskeletal fitness, including upper-body, middle-body and lower-body strength, endurance and power, and flexibility.

***Table 3:*** *Field-based fitness tests utilised to assess musculoskeletal fitness*

| **Fitness Test** | **Fitness test battery** | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | ALPHA | EUROFIT | FITNESSGRAM | PCPF | AAUTB | YMCAYFT | NYPFP | HRFT | Physical Best | NZFT | AAHPER | Bouge | CAHPER-FPT | PFTB | INDARES | PFAAT | PROESP | SLOfit | UNIFITTEST |
| ***Maximal isometric strength*** | | | | | | | | | | | | | | | | | | | |
| Handgrip strength | X | X | | | | | | | | | | | | | | X | | | |
| Trunk lift | | X | | X | | | | | | | | | | | | | | | |
| ***Endurance (upper body)*** | | | | | | | | | | | | | | | | | | | |
| Bent arm hang | | X | X | X | X | | X | | X | | X | | X | | | | | X | |
| Pull ups | | | X | X | | | X | | | | X | | | | | X | | | |
| Modified pull ups | | | X | | | X | X | | | | | | | | | | | | |
| Push ups | | | X | X | | | X | | | | | | | | X | | | | |
| Modified push ups | | | | | X | | X | | | | | | | | | | | | |
| Isometric push up | | | | | X | | | | | | | | | | | | | | |
| Parallel bar dips | | | | | | | X | | | | | | | | | | | | |
| ***Endurance (trunk)*** | | | | | | | | | | | | | | | | | | | |
| Sit ups | | X | | | X | | X | X | X | | X | | X | | | | X | X | X |
| Curl ups | | | X | X | | X | | | | X | | X | | | | | | | |
| Modified curl up | | | | | | | | | | | | | | | X | | | | |
| ***Endurance (lower body)*** | | | | | | | | | | | | | | | | | | | |
| Phantom chair | | | | | X | | | | | | | | | | | | | | |
| Chair squats | | | | | | | | | | | | | | | X | | | | |
| ***Explosive (upper body)*** | | | | | | | | | | | | | | | | | | | |
| Medicine ball throw | | | | | | | | | | X | | | | X | | | X | | X |
| Sand ball throw (0.25kg) | | | | | | | | | | X | | | | | | | | | |

| **Fitness Test** | **Fitness test battery** | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | ALPHA | EUROFIT | FITNESSGRAM | PCPF | AAUTB | YMCAYFT | NYPFP | HRFT | Physical Best | NZFT | AAHPER | Bouge | CAHPER-FPT | PFTB | INDARES | PFAAT | PROESP | SLOfit | UNIFITTEST |
| Shot put | | | | | | | | | | X | | | | | | | | | |
| Soft ball throw | | | | | | | | | | | X | | | | | X | | | |
| Cricket ball throw | | | | | | | | | | | | | | | X | | | | |
| Basketball throw | | | | | | | | | | | | X | | | | | | | |
| Tennis ball throw | | | | | | | | | | | | | | X | | | | | |
| ***Explosive (lower body)*** | | | | | | | | | | | | | | | | | | | |
| Standing broad jump | X | X | | | X | | X | | | X | X | X | X | X | | X | X | X | X |
| Vertical jump | | | | | | | | | | | | | | | | X | | | |
| ***Flexibility*** | | | | | | | | | | | | | | | | | | | |
| Sit and reach | | X | X | X | X | | | X | X | X | | X | | | | X | X | | |
| V-sit and reach | | | | | | | | | | | | | | | X | | | | |
| Back-saver sit and reach | X | | | | | | | | | | | | | | | | | | |
| Shoulder stretch | | | X | X | | | | | | | | X | | | X | | | | |
| Stand and reach | | | | | | | | | | | | | | | | X | | X | |

Safrit (1986) believed that for a test or a fitness test battery to be considered “good”, it should measure what it is supposed to measure and, therefore, demonstrate “good” validity. The field-based fitness test can be compared to a “gold standard” (criterion measure) to establish its criterion-related validity (Docherty, 1996). The most appropriate musculoskeletal fitness tests must be utilised within the fitness test battery to ensure an accurate measure of the fitness component it is purported to measure. Mayorga-Vega, et al. (2015) highlighted the importance of ensuring that the choice of physical fitness test should be based on the test’s validity and feasibility of implementation. There is strong evidence that the handgrip strength test is valid for assessing isometric muscular strength (Castro-Pinero, 2009). However, a systematic review found that, due to a limited number of studies, there was limited evidence that the standing long jump (SLJ) and vertical jump tests were valid for assessing explosive strength (Castro-Pinero, 2009). Castro-Pinero et al.., in a later study, indicate that the SLJ test is strongly associated with other tests that assess upper and lower body muscular strength. Based on these results, the SLJ test could be considered a general index of upper and lower body muscular fitness in youth.

Milliken et al. used the 1RM leg press as a gold standard and found that the SLJ test and the vertical jump test, along with body mass index (BMI), accounted for 44 and 41% of the variation in the 1RM leg press in children aged 7–12 years, respectively. Fernandez-Santos et al. (2015) demonstrated that the SLJ test was the lower-body muscular power test most closely associated with the criterion measure of the 1RM leg extension and seems to be the most valid field-based lower-body muscular power test.

Castro-Pinero (2010) found limited evidence that the bent arm hang, push-up, and pull-up tests are valid for assessing muscular endurance. The back-saver sit and reach test was moderately valid in assessing hamstring flexibility and low in assessing low back flexibility. Regarding musculoskeletal fitness, handgrip strength (HGS) is associated with overall muscular strength and endurance and is used as a screening tool in various branches of medicine and sports sciences (Trosclair et al., 2011; Buck et al., 2022). The Standing Long Jump (SLJ) was strongly associated with other lower-body muscular strength tests ($R^2$ = 0.83 – 0.86) and upper-body muscular strength tests ($R^2$ = 0.69 – 0.85). Therefore, The SLJ test is considered a general index of muscular fitness in youth that represents muscular strength, power, and local muscular endurance (Faigenbaum et al., 2020 and Castro-Pinero et al., 2010). In terms of feasibility and scalability within the school setting, both HGS and SLJ tests are time-efficient, cost-effective and easy to administer (Ruiz et al.., 2010; Castro-Pinero et al.., 2010). Cruz-León et al. (2025) found in their systematic review that the HGS and SLJ tests demonstrated strong evidence of feasibility.

Following the above literature review, the HGS and SLJ will be considered the most appropriate tests for musculoskeletal fitness in the primary school setting. As part of this study, they will undergo further testing to establish the typical measurement error and the smallest worthwhile change for each age group and sex.

## 2.3. Body Composition

Overweight and obesity in children and adolescents have become a global health problem with an increasing prevalence. Intra-abdominal adipose tissue accumulation increases the risk of developing insulin resistance, diabetes, and cardiovascular disease in adulthood (Simoni, 2023). Armstrong et al., 2006 showed that South African children have similar trends in obesity and overweight to developing countries. Strong evidence indicates that body composition in childhood and adolescence predicts cardiovascular risk factors such as carotid artery intima-media thickness and blood lipids (Ruiz et al., 2009; Juonala et al., 2006). There is also strong evidence to indicate that a high body mass index in childhood and adolescence increases the risk of death later in life (Ruiz, 2009). In a study by Garnet et al. (2007), boys and girls who were overweight or obese at 8 years of age were 7 times (Odds Ratio: 6.9; 95% CI 2.5 to 19.0; $p < 0.001$) as likely to have CVD risk clustering in adolescence compared with their peers who were not overweight or obese. They also showed that those children with increased central adiposity (measured by waist circumference) at 8 years old were four times (95% CI 3.6; 1.0 to 12.9; $p=0.061$) as likely to have cardiovascular disease risk clustering in adolescence than were children with a smaller waist circumference. In a study by Baker et al. (2007), the risk of any cardiovascular disease event (non-fatal or fatal) among adults was positively associated with BMI at 7 - 13 years of age for boys and 10 - 13 years of age for girls. Adjustment for birth weight strengthened the results.

Many authors have explored whether there is an association between back pain and body composition. Moderate evidence indicates no association between body composition (body mass index) and low back pain (Ruiz et al., 2009; Mikkelsson et al., 2006; Hestbaek et al., 2006). Body composition presents strong predictive evidence for being related to future disease and/or better health status, so it seems reasonable to include physical fitness tests in a battery of tests to be used in the school setting. Table 4 shows seven field-based tests used to assess body composition.

***Table 4:*** *Field-based fitness tests utilised to assess body composition*

| **Fitness Test** | **Fitness test battery** | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | ALPHA | EUROFIT | FITNESSGRAM | PCPF | AAUTB | YMCAYFT | NYPFP | HRFT | Physical Best | NZFT | AAHPER | Bouge | CAHPER-FPT | PFTB | INDARES | PFAAT | PROESP | SLOfit | UNIFITTEST |
| ***Anthropometric indices*** | | | | | | | | | | | | | | | | | | | |
| Height and weight | X | X | X | X | | X | | X | | X | | X | | | X | X | X | X | |
| Waist circumference | X | | | | | | | | | | | | | | | | X | | |
| Body mass index | X | | X | X | | | | | X | | | X | | | | X | X | | |
| Body fat % (skinfolds) | X | X | X | X | | X | | X | X | X | | | | | X | | | X | |
| Height to waist ratio | | | | | | | | | | | | | | | | | X | | |
| Wingspan | | | | | | | | | | | | | | | | | X | | |
| Bioelectrical impedance | | | X | | | | | | | | | | | | X | | | | |

*Table adapted/compiled from Ruiz et al., 2010; Marques et al., 2021*

Ensuring that the field-based test has good criterion-related validity and can accurately measure the fitness component it is alleged to measure is important (Safrit et al., 1986; Docherty et al., 1996). This is an essential aspect of its inclusion in a test battery (Ruiz et al., 2011). There is strong evidence that skinfold thickness (using Slaughter's equations) and body mass index (BMI) are good predictors of body fat. Slaughter's equations appear less accurate in obese children (Castro-Pinero et al., 2010). Here, BMI seems to be the best indicator, where Castro-Pinero et al. (2010) and Pietrobelli et al. (1998) report strong associations with total body fat ($R^2$ = 0.85 and 0.89 for boys and girls, respectively) and percent body fat ($R^2$ = 0.63 and 0.69 for boys and girls, respectively). However, its degree of accuracy does vary by the degree of body fatness, significantly improving at higher levels of body fat (Semiz et al., 2007; Moreno et al., 2006; Freedman et al., 2005). Strong evidence exists that waist circumference is a valid measure to estimate central fatness in children and adolescents ($r=0.92$, $P<0.001$) in both boys and girls (Taylor et al., 2000; Castro-Pinero et al., 2010).

Feasibility and scalability of fitness tests are worth considering when decisions are being made about which test/s to utilise (Espana-Romero et al., 2010; Domone et al., 2016). Espana-Romeno (2010) found that assessing skinfold thickness in the school setting was more difficult than 20mSRT, HGS and SLJ when utilising physical education teachers to administer the tests. Skinfold thickness testing was time-consuming and required many hours of experience for testers to develop an adequate level of proficiency in administering the tests. The other body composition measures of height, weight (BMI) and waist circumference were easier to administer from a practical and feasibility perspective. Body mass index z-score analysis is a method that can be used to assess children's weight relative to their peers (Martinez-Millana, 2018). A z-score of between −2 and +0.99 can be classified as underweight/normal weight. Z-scores from 1 to 1.99 is considered overweight and > 2 to 2.99 is considered obese (Monasor-Ortolá et al, 2021).

After reviewing the literature on body composition and its predictive validity, criterion validity, and feasibility, body mass index and waist circumference will be considered the most appropriate tests for body composition in primary school. As part of this study, they will undergo further testing to establish the typical measurement error and smallest worthwhile change for each age group and sex.

## 2.4. Motor Fitness

Motor fitness encompasses learning and performing motor skills, such as agility, balance, coordination, speed and power-related activities. Motor fitness is described in the literature by many authors as part of skill-related or performance-related fitness and not health-related fitness (Howley et al., 2001; Haga et al., 2008; Plowman and Meredith, 2013; Ortega et al., 2015). As per Table 5, several field-based fitness tests have been utilised to assess motor fitness.

***Table 5:*** *Field-based fitness tests utilised to assess motor fitness*

| Fitness Test | Fitness test battery | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | ALPHA | EUROFIT | FITNESSGRAM | PCPF | AAUTB | YMCAYFT | NYPFP | HRFT | Physical Best | NZFT | AAHPER | Bouge | CAHPER-FPT | PFTB | INDARES | PFAAT | PROESP | SLOfit | UNIFITTEST |
| ***Agility*** | | | | | | | | | | | | | | | | | | | |
| 10 x 5m shuttle run | | X | | | | | | | | | | X | | X | | | | | |
| 4 x 10m shuttle run | X | | | | X | | | | | | | | X | | X | | | | X |
| 4 x 30ft shuttle run | | | | | | | | | | | X | | | | | | | | |
| Climbing up wall bars | | | | | | | | | | | | | | X | | | | | |
| Side to side steps | | | | | | | | | | | | | | | | X | | | |
| Square test (4 x 4m) | | | | | | | | | | | | | | | | | X | | |
| Polygon backward | | | | | | | | | | | | | | | | | | X | |
| ***Speed*** | | | | | | | | | | | | | | | | | | | |
| 100m sprint | | | | | X | | | | | | | | X | | | | | | |
| 60m sprint | | | | | | | | | | | | | | | X | | | X | |
| 50m sprint | | | | | X | | | | | | | X | X | | | X | | | |
| 50-yard dash | | | | | | | | | | | X | | | | | | | | |
| 20m sprint | | | | | | | | | | | | | | X | | | X | | |
| ***Speed of limbs*** | | | | | | | | | | | | | | | | | | | |
| Plate tapping | | X | | | | | | | | | | | | | | | | X | |
| ***Balance*** | | | | | | | | | | | | | | | | | | | |
| Flamingo balance | | | X | | | | | | | | | | | | | | | | |

Ruiz et al. (2009) reported that motor fitness was not associated with total cholesterol, HDL cholesterol, or the ratio of both. There is inconclusive evidence indicating that motor fitness in childhood and adolescence is a predictor of cardiovascular disease risk factors later in life (Ruiz et al., 2009; Castro-Pinero et al., 2010). Although important, aspects of motor fitness do not seem as closely related to future disease and/or better health status when compared to cardiorespiratory fitness, musculoskeletal fitness and body composition. For this reason, many authors have not considered motor fitness tests ideal for inclusion in a battery of fitness tests for children (Cvejić et al., 2013). Therefore, motor fitness field-based tests will not be included for further evaluation of typical error of measurement and smallest worthwhile change.

## Summary

In conclusion, based on the review of the literature, the following physical fitness tests will be considered the most appropriate for use within the primary school setting. They will, as part of this study, undergo further testing to establish the typical error of measurement and the smallest worthwhile change for each age group and sex.

1). 20m shuttle run test

2). Hand grip strength

3). Standing long jump

4). Waist circumference

5). Body mass index

# CHAPTER 3: METHODS

## Participant Information

Two hundred and ninety volunteer learners from a South African Primary School were recruited for the study via convenience sampling methods. The sample, consisting of males (n = 156) and females (n = 134) ranging from 6 to 13 years of age, was divided into groups by sex (n = 2) and grade (n = 7) (i.e., a total of fourteen sub-groups).

## Study Design

Cross sectional study where each participant performed the testing protocol three times within a 7-day period, with a minimum of two days between each testing session. The tests were performed at the same time of day (within two hours) to limit variation because of circadian rhythm.

## Procedures

A standardised, 5-minute general warm-up consisting of jogging, dynamic stretches (arms swings, leg swings), squats and squat jumps were included before testing. The testing was set up to ensure participants were not fatigued when performing subsequent tests by alternating body composition and musculoskeletal testing stations and performing the cardiorespiratory fitness test last. The following tests were performed:

## Standing Height

Standing height (cm) was recorded as the height from the floor to the vertex of the head using a stadiometer. Participants stood barefoot with their arms hanging at their sides. The tester ensured that the participant's heels, buttocks, upper back and head were in contact with the wall. Two measurements were recorded and the mean of the two was used for analysis (accurate to 0.1 cm).

## Weight

Weight (in kilograms) was recorded using a Tanita BC-532 calibrated scale (Tanita Corporation, Tokyo, Japan). Participants were weighed wearing shorts and t-shirt and without shoes. They were instructed to stand in the centre of the scale with their weight evenly distributed between both feet and without support. Two measures were recorded and the mean of the two was used for analysis (accurate to 0.1 kg).

## Waist Circumference

Participants stood with their abdomen relaxed, arms at their sides and feet together. The administrator faced the child and placed the non-elastic tape measure around him/her, in a horizontal plane, at the level of the natural waist (the narrowest part of the torso) as seen from the anterior aspect. In some obese children, it may be difficult to identify a waist narrowing. In such cases, the smallest horizontal circumference was measured in the area between the superior iliac spine and the costal edge in the mid-axillary line. The measurement was taken at the end of a normal expiration while ensuring the tape was not compressing the skin and was recorded to the nearest 0.1cm.

## Hand Grip Strength

HGS was measured with a Jamar (5030J1) hydraulic hand dynamometer (Patterson Medical, Canada), set to handlebar position two (Gasior, 2018 and Molenaar, 2008). During the test, the dynamometer was held in line with the forearm, hanging down at the side with the elbow fully extended. The arm and hand holding the dynamometer were not allowed to touch the body. The participant squeezed the hand dynamometer gradually and continuously for at least 2 seconds and a maximum of 5 seconds (España-Romero, 2010).

Following a short rest, a second attempt was made, and the best result (of each hand) was recorded (accurate to 0.1 kg).

## Standing Long Jump

To start the standing long jump test, the participant stood with their feet shoulder-width apart and toes behind a line. They then jumped as far as possible, trying to land with their feet together. They had two attempts, the best of which was recorded for analysis. The participants were instructed to “stick the landing”. A further attempt was allowed if the participant fell backwards. The distance (cm) was measured from the take-off line to the point where the back of the heel nearest to the take-off line landed on the ground.

## 20m Shuttle Run Test

Participants ran between two lines 20 m apart while keeping pace with audio signals from a pre-recorded audio device. The initial speed was 8.5 $km.h^{-1}$, which is increased by 0.5 $km.h^{-1}$ per minute (one-minute equals one stage). They were instructed to run in a straight line, to pivot on completing a shuttle, and to pace themselves following the audio signals. The test finished when the child failed to reach the end lines concurrent with the audio signals on two consecutive occasions. The last completed half-stage was recorded as their score.

**Table 23: 20mSRT data for stages and distances**

| Level | Shuttle | Total Shuttles | Distance |
|---|---|---|---|
| 0 | 0 | 0 | 0 |
| 1 | 7 | 7 | 140 |
| 2 | 8 | 15 | 300 |
| 3 | 8 | 23 | 460 |
| 4 | 8 | 31 | 620 |
| 5 | 9 | 40 | 800 |
| 6 | 9 | 49 | 980 |
| 7 | 10 | 59 | 1180 |
| 8 | 10 | 69 | 1380 |
| 9 | 10 | 79 | 1580 |
| 10 | 11 | 90 | 1800 |
| 11 | 11 | 101 | 2020 |
| 12 | 12 | 113 | 2260 |
| 13 | 12 | 125 | 2500 |
| 14 | 13 | 138 | 2760 |
| 15 | 13 | 151 | 3020 |
| 16 | 13 | 164 | 3280 |
| 17 | 14 | 178 | 3560 |
| 18 | 14 | 192 | 3840 |
| 19 | 15 | 207 | 4140 |
| 20 | 15 | 222 | 4440 |

## Ethical Approval and Considerations

The study protocol was approved by the University of Cape Town Human Research Ethics Committee and was conducted in accordance with the Declaration of Helsinki (World Medical Association, 2013). Permission to conduct the study was obtained from the Western Cape Department of Education (DoE) and the principal of the school.

The nature and scope, as well as the risks and benefits of the study, were explained to the school children and their parents via letters (Appendix 1 and 2). Informed written consent was obtained from the parents/guardian of each participant (Appendix 3). Written assent was also obtained from each participant (Appendix 4). The parents of all children involved completed a physical activity readiness questionnaire

(Appendix 5). Participants had to be full-time learners at the school to be included in the study. Participants were excluded from the research study if they had an injury or any medical condition that precluded them from participating in moderate to strenuous physical activity at the time of the study. Participants were also excluded if their parents did not consent or the child did not give his/her assent. All data were kept confidential.

## Statistical Analysis

SPSS Software (version 26) was used to analyse the data. All data are expressed as mean ± standard deviation and statistical significance was accepted at $P < 0.050$. A paired t-test was used to determine differences between left and right HGS. The relationships between HGS, body mass, and height were determined with a Pearson's correlation using GraphPad Prism version 8.0.0 for Mac, GraphPad Software, San Diego, California USA, www.graphpad.com.

The typical error of measurement (TE) and coefficient of variation ($CV_{TE}$) were calculated using the spreadsheet '*Consecutive pairwise analysis of trials for reliability*' (downloaded from www.sportsci.org). Reliability measures are expressed with 90% confidence intervals (90% CI). Between-day differences were examined for statistical significance using a repeated measures analysis of variance. Significant differences were further analysed with a Tukey's *post-hoc* test. Cohen's effect sizes (*d*) were calculated to investigate the magnitude of between-trial differences. Effect sizes were interpreted as < 0.2 = *trivial*, ≥ 0.2 to < 0.5 = *small*, ≥ 0.5 to < 0.8 = *moderate* and ≥ 0.8 = *large* (Cohen, 1988).

The sensitivity of each of the tests was determined by comparing the smallest worthwhile change (SWC) to the TE in absolute units. The SWC representing a *small, medium* and *large* change was calculated as 0.2, 0.5 and 0.8 × between-subject SD, respectively. TE < SWC was defined as 'good', TE=SWC as 'OK', and TE > SWC as 'marginal' (Hopkins, 2004).

# CHAPTER 4: RESULTS

Participants' age, body mass, standing height and body mass index are shown in Table 6. The variables are arranged according to grade and sex. Values for body mass, standing height and body mass index are similar in both sexes with no significant differences.

***Table 6:*** *Age, body mass, height and body mass index for males (M) and females (F). Data expressed as mean ± SD (sample size)*

| Grade | Sex | Age (years) | Body mass (kg) | Height (cm) | BMI (kg/m$^2$) |
|---|---|---|---|---|---|
| 1 | M | 7.3 ± 0.4 (28) | 23.9 ± 5.5 (28) | 122.8 ± 4.9 (28) | 15.7 ± 2.5 (28) |
| 1 | F | 7.1 ± 0.4 (22) | 26.1 ± 4.3 (22) | 124.5 ± 5.0 (22) | 16.8 ± 1.9 (22) |
| *Total* | | *7.2 ± 0.4 (50)* | *24.9 ± 5.0 (50)* | *123.6 ± 5.0 (50)* | *16.2 ± 2.3 (50)* |
| 2 | M | 8.1 ± 0.3 (27) | 27.6 ± 4.5 (27) | 131.3 ± 5.3 (27) | 15.9 ± 1.8 (27) |
| 2 | F | 8.2 ± 0.4 (18) | 32.6 ± 7.1 (18) | 133.4 ± 6.1 (18) | 18.2 ± 2.8 (18) |
| *Total* | | *8.2 ± 0.3 (45)* | *29.6 ± 6.1 (45)* | *132.2 ± 5.7 (45)* | *16.8 ± 2.5 (45)* |
| 3 | M | 9.1 ± 0.4 (24) | 33.3 ± 9.5 (24) | 136.0 ± 6.8 (24) | 17.8 ± 3.5 (24) |
| 3 | F | 9.1 ± 0.3 (23) | 34.0 ± 10.1 (23) | 137.6 ± 8.3 (23) | 17.7 ± 3.5 (23) |
| *Total* | | *9.1 ± 0.3 (47)* | *33.7 ± 9.7 (47)* | *136.8 ± 7.5 (47)* | *17.7 ± 3.5 (47)* |
| 4 | M | 10.2 ± 0.3 (22) | 38.0 ± 11.9 (20) | 142.4 ± 7.1 (20) | 18.4 ± 4.4 (20) |
| 4 | F | 10.2 ± 0.4 (15) | 35.2 ± 9.5 (15) | 140.3 ± 8.6 (15) | 17.6 ± 3.6 (15) |
| *Total* | | *10.2 ± 0.4 (37)* | *36.8 ± 10.9 (35)* | *141.5 ± 7.7 (35)* | *18.1 ± 4.1 (35)* |
| 5 | M | 11.1 ± 0.3 (23) | 40.9 ± 10.4 (22) | 148.1 ± 6.7 (22) | 18.5 ± 3.6 (22) |
| 5 | F | 11.1 ± 0.3 (19) | 43.1 ± 9.0 (18) | 151.3 ± 8.3 (19) | 18.7 ± 2.6 (18) |
| *Total* | | *11.1 ± 0.3 (42)* | *41.9 ± 9.8 (40)* | *150.0 ± 7.6 (41)* | *18.6 ± 3.2 (40)* |
| 6 | M | 12.1 ± 0.3 (23) | 42.7 ± 6.6 (22) | 153.1 ± 7.6 (22) | 18.2 ± 2.3 (22) |
| 6 | F | 12.1 ± 0.3 (14) | 42.8 ± 8.3 (14) | 153.5 ± 6.9 (14) | 17.9 ± 2.3 (14) |
| *Total* | | *12.1 ± 0.3 (37)* | *42.7 ± 7.2 (36)* | *153.3 ± 7.3 (36)* | *18.1 ± 2.3 (36)* |
| 7 | M | 13.2 ± 0.4 (13) | 49.1 ± 14.5 (13) | 158.2 ± 9.7 (13) | 19.3 ± 3.8 (13) |
| 7 | F | 13.1 ± 0.4 (23) | 47.8 ± 14.1 (23) | 157.1 ± 7.6 (23) | 19.1 ± 4.0 (23) |
| *Total* | | *13.1 ± 0.4 (36)* | *48.3 ± 14.0 (36)* | *157.5 ± 8.3 (36)* | *19.2 ± 3.9 (36)* |
| All | M | 9.8 ± 1.9 (160) | 34.9 ± 11.8 (156) | 139.6 ± 13.2 (156) | 17.5 ± 3.3 (156) |
| | F | 10.1 ± 2.1 (134) | 37.2 ± 11.8 (133) | 142.1 ± 13.5 (134) | 18.0 ± 3.1 (133) |
| *Total* | | 9.9 ± 2.0 (294) | 36.0 ± 11.9 (289) | 140.8 ± 13.4 (290) | 17.7 ± 3.2 (289) |

## Body Mass Index (BMI)

Table 7 shows the BMI for males and females from Grades 1 to 7. BMI increased with increasing age in both males and females, except in Grade 6, where there was a slight decrease from Grades 4 and 5 in males and Grade 5 in females. The mean and standard deviation for BMI on the three testing days was 17.7 kg/m$^2$ (± 3.2 kg/m$^2$), and there were no significant differences between the three testing days. The mean of the three days was therefore used for subsequent calculations. The results of an analysis of variance (ANOVA) are also shown in Table 7. The main effect of grade (age) was significantly different (P < 0.001).

***Table 7:*** *Body Mass Index (kg/m2) of children from Grades 1 to 7 (n=289). Values are expressed as mean ± SD (sample size)*

| **Grade**[a] | **Female** | **N** | **Male** | **n** | **Total**[b] | **n** |
|---|---|---|---|---|---|---|
| 1 | 16.8 ± 1.9 | 22 | 15.7 ± 2.5 | 28 | *16.2 ± 2.3* | 50 |
| 2 | 18.2 ± 2.8 | 18 | 15.9 ± 1.8 | 27 | *16.8 ± 2.5* | 45 |
| 3 | 17.7 ± 3.5 | 23 | 17.8 ± 3.5 | 24 | *17.7 ± 3.5* | 47 |
| 4 | 17.6 ± 3.6 | 15 | 18.4 ± 4.4 | 20 | *18.1 ± 4.1* | 35 |
| 5 | 18.7 ± 2.6 | 18 | 18.5 ± 3.6 | 22 | *18.6 ± 3.2* | 40 |
| 6 | 17.9 ± 2.3 | 14 | 18.2 ± 2.3 | 22 | *18.1 ± 2.3* | 36 |
| 7 | 19.1 ± 4.0 | 23 | 19.3 ± 3.8 | 13 | *19.2 ± 3.9* | 36 |
| *Total* | 18.0 ± 3.1 | *133* | 17.5 ± 3.3 | *156* | 17.7 ± 3.2 | *289* |

| | | |
|---|---|---|
| [a]*Main effect of Grade* | P < 0.001 | $F_{6, 275} = 4.1$ |
| 1 vs 5, 7 | *P < 0.005* | |
| 2 vs 7 | *P < 0.015* | |
| [b]*Main effect of Gender* | P = 0.396 | $F_{1, 275} = 0.7$ |
| *Interaction Grade x Gender* | P = 0.338 | $F_{6, 275} = 1.1$ |

The TE and SWC arranged according to grade, and sex are shown in Table 8. The ability of BMI measurements to detect *small*, *medium* and *large* changes is summarised in this table. All the grades and sexes showed a SWC that had a 'good' ability to detect *small, medium* and *large* changes as the SWC was greater than the TE for each of these grades and sexes.

***Table 8:** Ability of BMI Measurement to detect Small, Medium and Large Changes using Typical Error of Measurement (kg/m2) and Smallest Worthwhile Change (kg/m2) for BMI (kg/m2) of children from Grades 1 to 7 (n=289). Data expressed as mean (90% CI).*

| | | TE | SWC | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| **Sex** | **Grade** | | Small | Test Rating | Medium | Test Rating | Large | Test Rating | N |
| Females | 1 | 0.2 (0.2-0.2) | 0.4 | Good | 0.9 | Good | 1.5 | Good | 22 |
| | 2 | 0.2 (0.2-0.3) | 0.6 | Good | 1.4 | Good | 2.2 | Good | 18 |
| | 3 | 0.1 (0.1-0.2) | 0.7 | Good | 1.8 | Good | 2.9 | Good | 23 |
| | 4 | 0.2 (0.1-0.2) | 0.7 | Good | 1.7 | Good | 2.8 | Good | 15 |
| | 5 | 0.1 (0.1-0.2) | 0.5 | Good | 1.3 | Good | 2.0 | Good | 18 |
| | 6 | 0.1 (0.1-0.2) | 0.5 | Good | 1.2 | Good | 1.9 | Good | 14 |
| | 7 | 0.2 (0.1-0.2) | 0.8 | Good | 2.0 | Good | 3.2 | Good | 23 |
| | *Total* | *0.2 (0.2-0.2)* | *0.6* | *Good* | *1.6* | *Good* | *2.5* | *Good* | *133* |
| Males | 1 | 0.2 (0.2-0.3) | 0.5 | Good | 1.3 | Good | 2.1 | Good | 28 |
| | 2 | 0.2 (0.1-0.2) | 0.4 | Good | 0.9 | Good | 1.4 | Good | 27 |
| | 3 | 0.2 (0.2-0.2) | 0.7 | Good | 1.8 | Good | 2.8 | Good | 24 |
| | 4 | 0.1 (0.1-0.2) | 0.7 | Good | 1.8 | Good | 2.9 | Good | 20 |
| | 5 | 0.1 (0.1-0.2) | 0.7 | Good | 1.8 | Good | 3.0 | Good | 22 |
| | 6 | 0.1 (0.1-0.2) | 0.5 | Good | 1.2 | Good | 1.9 | Good | 22 |
| | 7 | 0.2 (0.2-0.3) | 0.7 | Good | 1.7 | Good | 2.7 | Good | 13 |
| | *Total* | *0.2 (0.2-0.2)* | *0.6* | *Good* | *1.6* | *Good* | *2.5* | *Good* | *156* |
| *All* | *1-7* | *0.2 (0.2-0.2)* | *0.6* | *Good* | *1.6* | *Good* | *2.5* | *Good* | *289* |

TE = Typical Error of Measurement; SWC = Smallest Worthwhile Change; Small = 0.2 × between subject SD; Medium = 0.5 × between subject SD; Large = 0.8 × between subject SD; TEM<SWC (Good); TEM=SWC (OK); TEM>SWC (Marginal); $kg/m^2$ = kilograms per metre squared; CI = Confidence Interval

## Waist Circumference

For the three testing days, the mean and standard deviation of waist circumference was 62.9 cm (± 7.8 cm), and there were no significant differences between the testing days. The mean of the three days was therefore used for subsequent calculations. Table 9 shows that waist circumference increased with increasing age in both males and females, except for Grade 6, where there was a slight decrease from Grade 5 before increasing again in Grade 7. The results of an analysis of variance (ANOVA) are also shown in Table 9 where the main effect of grade (age) was significantly different ($P < 0.001$).

***Table 9:*** *Waist Circumference (cm) of children from Grades 1 to 7 (n=290). Values are expressed as mean ± SD.*

| **Grade**[a] | **Female** | **n** | **Male** | **N** | **Total**[b] | **n** |
|---|---|---|---|---|---|---|
| 1 | 58.6 ± 4.5 | 22 | 57.2 ± 5.1 | 28 | *57.8 ± 4.8* | 50 |
| 2 | 62.6 ± 6.6 | 18 | 58.5 ± 4.3 | 27 | *60.1 ± 5.6* | 45 |
| 3 | 62.8 ± 8.6 | 23 | 62.9 ± 8.7 | 24 | *62.8 ± 8.6* | 47 |
| 4 | 63.7 ± 9.2 | 15 | 65.6 ± 10.6 | 20 | *64.8 ± 10.0* | 35 |
| 5 | 64.7 ± 6.5 | 19 | 66.2 ± 8.1 | 22 | *65.5 ± 7.3* | 41 |
| 6 | 62.7 ± 3.8 | 14 | 65.9 ± 3.5 | 22 | *64.6 ± 3.9* | 36 |
| 7 | 66.1 ± 9.4 | 23 | 68.5 ± 7.4 | 13 | *66.9 ± 8.7* | 36 |
| *Total* | 63.0 ± 7.6 | *134* | 62.8 ± 7.9 | *156* | 62.9 ± 7.8 | *290* |

| | | |
|---|---|---|
| [a]*Main effect of Grade* | P < 0.001 ($3.4584^{-8}$) | $F_{6, 276} = 8.2$ |
| 1 vs 3, 4, 5, 6, 7 | *P < 0.012* | |
| 2 vs 5, 7 | *P < 0.011* | |
| [b]*Main effect of Gender* | P = 0.531 | $F_{1, 276} = 0.4$ |
| *Interaction Grade x Gender* | P = 0.283 | $F_{6, 276} = 1.2$ |

The TE and SWC for waist circumference arranged according to grade and sex is shown in Table 10, where the ability of waist circumference measurements to detect *small*, *medium* and *large* changes is summarised. All the grades and sexes showed a SWC with a 'good' ability to detect *medium* and *large* changes. The ability to detect *small* changes is 'marginal' in some grades across both sexes due to the TE being greater than the *small* SWC.

***Table 10:*** *Ability of Waist Circumference Test to detect Small, Medium and Large Changes using Typical Error of Measurement (cm) and Smallest Worthwhile Change (cm) for WC (cm) of children from Grades 1 to 7 (n=290). Data expressed as mean (90% CI).*

| | | TE | SWC | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| **Sex** | **Grade** | | Small | Test Rating | Medium | Test Rating | Large | Test Rating | N |
| Females | 1 | 1.9 (1.6-2.3) | 1.0 | Marginal | 2.5 | Good | 4.0 | Good | 22 |
| | 2 | 1.3 (1.1-1.7) | 1.4 | Good | 3.4 | Good | 5.5 | Good | 18 |
| | 3 | 1.1 (1.0-1.4) | 1.8 | Good | 4.4 | Good | 7.0 | Good | 23 |
| | 4 | 1.5 (1.3-2.1) | 1.8 | Good | 4.4 | Good | 7.0 | Good | 15 |
| | 5 | 1.0 (0.9-1.3) | 1.3 | Good | 3.3 | Good | 5.2 | Good | 19 |
| | 6 | 0.9 (0.7-1.1) | 0.8 | Marginal | 1.9 | Good | 3.1 | Good | 14 |
| | 7 | 1.1 (0.9-1.3) | 1.9 | Good | 4.7 | Good | 7.5 | Good | 23 |
| | *Total* | *1.4 (1.4-1.6)* | *1.5* | *Good* | *3.8* | *Good* | *6.1* | *Good* | *134* |
| Males | 1 | 1.7 (1.5-2.0) | 1.1 | Marginal | 2.7 | Good | 4.2 | Good | 28 |
| | 2 | 1.2 (1.0-1.4) | 0.8 | Marginal | 2.1 | Good | 3.3 | Good | 27 |
| | 3 | 1.2 (1.0-1.4) | 1.8 | Good | 4.4 | Good | 7.0 | Good | 24 |
| | 4 | 2.2 (1.8-2.7) | 1.8 | Marginal | 4.5 | Good | 7.2 | Good | 20 |
| | 5 | 1.5 (1.3-1.8) | 1.7 | Good | 4.1 | Good | 6.6 | Good | 22 |
| | 6 | 1.2 (1.0-1.4) | 0.8 | Marginal | 1.9 | Good | 3.0 | Good | 22 |
| | 7 | 1.6 (1.3-2.3) | 1.4 | Marginal | 3.4 | Good | 5.5 | Good | 13 |
| | *Total* | *1.5 (1.4-1.6)* | *1.5* | *Good* | *3.8* | *Good* | *6.1* | *Good* | *156* |
| *All* | *1-7* | *1.5 (1.4-1.6)* | *1.5* | *Good* | *3.8* | *Good* | *6.1* | *Good* | *290* |

TE = Typical Error of Measurement; SWC = Smallest Worthwhile Change; Small = 0.2 × between subject SD; Medium = 0.5 × between subject SD; Large = 0.8 × between subject SD; TE<SWC (Good); TE=SWC (OK); TE>SWC (Marginal); cm = centimetres; CI = Confidence Interval

## Hand Grip Strength (HGS)

The maximum average difference in HGS between testing days for the left and right hand was 0.34 kg and 0.43 kg, respectively. Although the differences were significant (left; $F_{2, 460} = 4.4$, p = 0.014) and (right; $F_{2, 473} = 8.5$; p = 0.001), the magnitude of the differences were much smaller than the TE, and therefore we used the average of the three days for subsequent calculations. Table 11 shows the hand grip strength for males and females from Grades 1 to 7. The grip strength increased in both hands with increasing age in both sexes. The main effect of sex was not significantly different. Also, the interaction between grade and sex was not significant. We observed a significant difference between left- and right-hand grip strength values (P < 0.001). Therefore, mean hand grip strength was utilised during further analysis. The results of an analysis of variance (ANOVA) are also shown in Table 11.

***Table 11:*** *Hand Grip Strength (HGS Left and HGS Right) for males (M) and females (F) from Grades 1 to 7. Data expressed (in kilograms) as mean ± SD (sample size).*

| **Grade**[a] | **Sex** | **HGS Left (Mean)** | **HGS Right (Mean)** | **HGS (Mean)** |
|---|---|---|---|---|
| 1 | M | 10.1 ± 2.4 (28) | 10.8 ± 2.3 (28) | 10.4 ± 2.3 (28) |
| 1 | F | 10.2 ± 2.2 (22) | 10.8 ± 2.4 (22) | 10.5 ± 2.3 (22) |
| *Total* | | *10.1 ± 2.3 (50)* | *10.8 ± 2.3 (50)* | *10.4 ± 2.2 (50)* |
| 2 | M | 12.9 ± 2.1 (27) | 13.5 ± 2.4 (27) | 13.2 ± 2.2 (27) |
| 2 | F | 13.3 ± 2.6 (18) | 13.7 ± 2.6 (18) | 13.5 ± 2.6 (18) |
| *Total* | | *13.1 ± 2.3 (45)* | *13.6 ± 2.5 (45)* | *13.3 ± 2.3 (45)* |
| 3 | M | 14.7 ± 3.0 (24) | 15.4 ± 3.1 (24) | 15.1 ± 2.9 (24) |
| 3 | F | 14.1 ± 3.4 (23) | 15.1 ± 3.3 (23) | 14.6 ± 3.3 (23) |
| *Total* | | *14.4 ± 3.2 (47)* | *15.3 ± 3.2 (47)* | *14.9 ± 3.1 (47)* |
| 4 | M | 17.0 ± 3.0 (20) | 17.5 ± 2.8 (20) | 17.2 ± 2.8 (20) |
| 4 | F | 15.5 ± 3.8 (15) | 15.7 ± 3.5 (15) | 15.6 ± 3.5 (15) |
| *Total* | | *16.4 ± 3.4 (35)* | *16.7 ± 3.2 (35)* | *16.5 ± 3.2 (35)* |
| 5 | M | 19.3 ± 4.2 (22) | 20.1 ± 3.9 (22) | 19.7 ± 4.0 (22) |
| 5 | F | 20.4 ± 3.3 (19) | 21.6 ± 4.0 (19) | 21.0 ± 3.6 (19) |
| *Total* | | *19.9 ± 3.8 (41)* | *20.8 ± 4.0 (41)* | *20.3 ± 3.8 (41)* |
| 6 | M | 22.7 ± 5.6 (22) | 23.4 ± 6.2 (22) | 23.0 ± 5.7 (22) |
| 6 | F | 21.3 ± 5.6 (14) | 22.1 ± 6.1 (14) | 21.7 ± 5.8 (14) |
| *Total* | | *22.1 ± 5.5 (36)* | *22.9 ± 6.1 (36)* | *22.5 ± 5.7 (36)* |
| 7 | M | 27.7 ± 7.9 (13) | 27.9 ± 6.5 (13) | 27.8 ± 7.1 (13) |
| 7 | F | 25.2 ± 5.0 (23) | 26.6 ± 4.5 (23) | 25.9 ± 4.7 (23) |
| *Total* | | *26.1 ± 6.2 (36)* | *27.0 ± 5.3 (36)* | *26.6 ± 5.7 (36)* |
| All | M | 16.7 ± 6.5 (156) | 17.4 ± 6.4 (156) | 17.0 ± 6.4 (156) |
| | F | 17.1 ± 6.3 (134) | 17.9 ± 6.6 (134) | 17.5 ± 6.4 (134) |
| *Total* | | 16.9 ± 6.4 (290) | 17.6 ± 6.5 (290) | 17.2 ± 6.4 (290) |

[a]*Main effect of Grade* P < 0.001 $F_{6, 276} = 86.9$

Grade 1 vs 2, 3, 4, 5, 6, 7 *P < 0.005*

Grade 2 vs 4, 5, 6, 7 *P < 0.004*

| | | |
|---|---|---|
| Grade 3 vs 5, 6, 7 | *$P < 0.001$* | |
| Grade 4 vs 5, 6, 7 | *$P < 0.001$* | |
| Grade 5 vs 7 | *$P < 0.001$* | |
| [b]*Main effect of Sex* | $P = 0.253$ | $F_{1, 276} = 1.3$ |
| *Interaction Grade x Sex* | $P = 0.514$ | $F_{6, 276} = 0.8$ |
| *HGS left vs. HGS right* | $P < 0.00001$ | |

The TE and SWC arranged according to grade and sex are shown in Table 12. There was no significant difference between the TE and SWC of the left and right hands. Table 12, therefore, shows a summary of the ability of the handgrip strength test to detect *small, medium* and *large* changes for the right hand. The SWC to detect *small* changes in HGS in each of the grades (and for both sexes) was less than the TE, showing a test rating of 'marginal'. The SWC to detect *medium* and *large* changes in HGS was greater than the TE for each of the grades and sexes, showing a test rating of 'good'.

***Table 12:*** *Ability of the Hand Grip Strength Test to detect Small, Medium and Large Changes using Typical Error of Measurement (kg) and Smallest Worthwhile Change (kg) for Right HGS (kg) of children from Grades 1 to 7 (n=290). Data expressed as mean (90% CI).*

| | | **TE** | **SWC** | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| **Sex** | **Grade** | | Small | Test Rating | Medium | Test Rating | Large | Test Rating | N |
| Females | 1 | 0.7 (0.6-0.9) | 0.5 | Marginal | 1.2 | Good | 1.9 | Good | 22 |
| | 2 | 1.4 (1.1-1.7) | 0.6 | Marginal | 1.4 | Good | 2.3 | Good | 18 |
| | 3 | 1.0 (0.8-1.2) | 0.7 | Marginal | 1.7 | Good | 2.7 | Good | 23 |
| | 4 | 1.1 (0.9-1.5) | 0.8 | Marginal | 1.9 | Good | 3.1 | Good | 15 |
| | 5 | 1.2 (1.1-1.5) | 0.9 | Marginal | 2.2 | Good | 3.5 | Good | 19 |
| | 6 | 1.4 (1.1-1.8) | 1.3 | Marginal | 3.3 | Good | 5.3 | Good | 14 |
| | 7 | 1.2 (1.0-1.4) | 0.9 | Marginal | 2.2 | Good | 3.6 | Good | 23 |
| | *Total* | *1.2 (1.1-1.3)* | *1.4* | *Good* | *3.4* | *Good* | *5.5* | *Good* | *134* |
| Males | 1 | 0.9 (0.8-1.1) | 0.5 | Marginal | 1.2 | Good | 1.9 | Good | 28 |
| | 2 | 1.2 (1.0-1.4) | 0.5 | Marginal | 1.2 | Good | 2.0 | Good | 27 |
| | 3 | 1.4 (1.2-1.7) | 0.7 | Marginal | 1.7 | Good | 2.7 | Good | 24 |
| | 4 | 1.5 (1.3-1.9) | 0.6 | Marginal | 1.5 | Good | 2.4 | Good | 20 |
| | 5 | 1.1 (1.0-1.4) | 0.8 | Marginal | 2.0 | Good | 3.2 | Good | 22 |
| | 6 | 1.5 (1.2-1.8) | 1.3 | Marginal | 3.2 | Good | 5.1 | Good | 22 |
| | 7 | 1.5 (1.2-2.1) | 1.3 | Marginal | 3.2 | Good | 5.1 | Good | 13 |
| | *Total* | *1.3 (1.3-1.4)* | *1.3* | *Marginal* | *3.3* | *Good* | *5.2* | *Good* | *156* |
| *All* | *1-7* | *1.3 (1.2-1.3)* | *1.3* | *Good* | *3.3* | *Good* | *5.3* | *Good* | *290* |

TE = typical error of measurement; SWC = smallest worthwhile change; Small = 0.2 × between subject SD; Medium = 0.5 × between subject SD; Large = 0.8 × between subject SD; TE<SWC (Good); TE=SWC (OK); TE>SWC (Marginal); kg = kilograms; CI = confidence interval

A significant relationship ($P < 0.0001$) existed between HGS and body mass (females: $r = 0.79$, 95% confidence interval 0.72 – 0.85, n = 133 and males: $r = 0.78$, 95% confidence interval 0.72 – 0.84, n = 156). A significant relationship ($P < 0.0001$) also existed between HGS and height (females: $r = 0.89$, 95% confidence interval 0.85 – 0.92, n = 134 and males: $r = 0.90$, 95% confidence interval 0.87 – 0.93, n = 156).

## Standing Long Jump (SLJ)

The mean and standard deviation for the SLJ on the three testing days were 132.7 cm (± 26.6 cm), 132.6 cm (± 25.5 cm) and 130.4 cm (± 28.3 cm). The third testing day's results were significantly lower than the first (P<0.023) and second (P<0.015) days' results. The small change (2.3 cm) test 1 vs. test 3 is less than the TE and the *small* SWC for SLJ. Based on these findings we used the first day's results for subsequent calculations.

Table 13 shows the SLJ for males and females from Grades 1 to 7. The SLJ increased in both females and males with increasing age, except for the Grade 3 males, where there was a slight decrease from Grade 2 to 3 before increasing again in Grade 4. The results of an analysis of variance (ANOVA) are also shown in Table 13. The main effect of sex was significantly different (P<0.001). The interaction between grade and sex was also significantly different for Grade 2 (P<0.008), Grade 6 (P<0.001) and Grade 7 (P<0.031).

***Table 13:*** *Standing Long Jump (cm) of children from Grades 1 to 7 (n=264). Values are expressed as mean ± SD.*

| **Grade[a]** | **Female** | **N** | **Male** | **n** | **Total[b]** | **n** |
|---|---|---|---|---|---|---|
| 1 | 105.8 ± 16.9 | 20 | 105.1 ± 19.1 | 26 | *105.4 ± 18.0* | *46* |
| 2 | 115.3 ± 14.7 | 16 | 131.9 ± 16.5 | 25 | *125.4 ± 17.7* | *41* |
| 3 | 121.9 ± 22.7 | 21 | 125.3 ± 15.6 | 21 | *123.6 ± 19.3* | 42 |
| 4 | 132.7 ± 17.0 | 13 | 138.2 ± 20.7 | 17 | *135.8 ± 19.1* | 30 |
| 5 | 142.9 ± 17.6 | 18 | 145.2 ± 18.9 | 20 | *144.1 ± 18.1* | 38 |
| 6 | 135.7 ± 20.4 | 12 | 161.7 ± 26.0 | 21 | *152.2 ± 26.9* | 33 |
| 7 | 153.9 ± 17.0 | 22 | 169.1 ± 28.0 | 12 | *159.3 ± 22.3* | 34 |
| *Total* | *129.8 ± 22.9* | *122* | *136.2 ± 28.3* | *142* | *133.2 ± 26.6* | *264* |

| | | |
|---|---|---|
| [a]*Main effect of Grade* | P < 0.001 ($1.8649^{-30}$) | $F_{6, 250} = 34.8$ |
| 1 vs 2, 3, 4, 5, 6, 7 | *P < 0.001* | |
| 2 vs 5, 6, 7 | *P < 0.001* | |
| 3 vs 5, 6, 7 | *P < 0.001* | |
| 4 vs 6, 7 | *P < 0.017* | |
| 5 vs 7 | *P < 0.019* | |
| [b]*Main effect of Sex* | P < 0.001 | $F_{1, 250} = 15.8$ |
| Male vs Female | *P < 0.001* | |
| *Interaction Grade x Sex* | P < 0.045 | $F_{6, 250} = 2.2$ |
| Grade 2 | *P < 0.008* | |
| Grade 6 | *P < 0.001* | |
| Grade 7 | *P < 0.031* | |

The TE and SWC arranged according to grade, and sex are shown in Table 14. The ability of the standing long jump test to detect *small, medium* and *large* changes is summarised in this table. The SWC to detect *small* changes in SLJ in each of the grades (and for both sexes) was less than the TE, showing a test rating of 'marginal'. The SWC to detect *medium* changes for Grade 1 females and Grade 2 males also showed a test rating of 'marginal'. The rest of the grades and sexes showed a SWC with a 'good' ability to detect *medium* and *large* changes in SLJ as the SWC was greater than the TE for each of these grades and sexes (Table 14).

***Table 14:*** *Ability of Standing Long Test to detect Small, Medium and Large Changes using Typical Error of Measurement (cm) and Smallest Worthwhile Change (cm) for SLJ (cm) of children from Grades 1 to 7 (n=289). Data expressed as mean (90% CI).*

| | | **TE** | **SWC** | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| **Sex** | **Grade** | | Small | Test Rating | Medium | Test Rating | Large | Test Rating | N |
| Females | 1 | 9.6 (8.2-11.7) | 3.2 | Marginal | 8.0 | Marginal | 12.7 | Good | 22 |
| | 2 | 8.4 (7.1-10.6) | 3.4 | Marginal | 8.4 | Good | 13.4 | Good | 17 |
| | 3 | 8.5 (7.3-10.4) | 4.0 | Marginal | 10.0 | Good | 16.0 | Good | 23 |
| | 4 | 8.3 (6.8-11.1) | 3.3 | Marginal | 8.3 | Good | 13.3 | Good | 15 |
| | 5 | 8.2 (6.9-10.3) | 3.7 | Marginal | 9.1 | Good | 14.6 | Good | 19 |
| | 6 | 5.4 (4.4-7.1) | 3.9 | Marginal | 9.9 | Good | 15.8 | Good | 14 |
| | 7 | 7.4 (6.4-8.9) | 3.5 | Marginal | 8.8 | Good | 14.1 | Good | 23 |
| | *Total* | *8.6 (8.1-9.3)* | *4.9* | *Marginal* | *12.2* | *Good* | *19.4* | *Good* | *133* |
| Males | 1 | 9.0 (7.8-10.6) | 4.1 | Marginal | 10.3 | Good | 16.5 | Good | 28 |
| | 2 | 9.9 (8.6-11.8) | 3.0 | Marginal | 7.5 | Marginal | 12.0 | Good | 27 |
| | 3 | 8.0 (6.9-9.8) | 3.4 | Marginal | 8.4 | Good | 13.4 | Good | 24 |
| | 4 | 8.4 (7.1-10.5) | 4.4 | Marginal | 11.0 | Good | 17.7 | Good | 20 |
| | 5 | 8.3 (7.1-10.1) | 4.2 | Marginal | 10.5 | Good | 16.8 | Good | 22 |
| | 6 | 8.2 (7.1-10.1) | 4.9 | Marginal | 12.3 | Good | 19.7 | Good | 22 |
| | 7 | 6.8 (5.6-9.5) | 5.3 | Marginal | 13.4 | Good | 21.4 | Good | 13 |
| | *Total* | *8.9 (8.4-9.6)* | *5.8* | *Marginal* | *14.4* | *Good* | *23.0* | *Good* | *156* |
| *All* | *1-7* | *8.8 (8.4-9.2)* | *5.4* | *Marginal* | *13.5* | *Good* | *21.5* | *Good* | *289* |

TE = Typical Error of Measurement; SWC = Smallest Worthwhile Change; Small = 0.2 × between subject SD; Medium = 0.5 × between subject SD; Large = 0.8 × between subject SD; TE<SWC (Good); TE=SWC (OK); TE>SWC (Marginal); kg = kilograms; CI = Confidence Interval

## 20m Shuttle Run Test (20mSRT)

Table 15 shows the 20mSRT from Grades 1 to 7 for males and females. The 20mSRT increased in females with increasing age, except for grade 7 where there was a slight decrease from grade 6 to 7. The males also increased with increasing age except for a slight decrease from Grade 2 to 3 before increasing again. There was a similar decrease from Grades 6 to 7. The results of an analysis of variance (ANOVA) are also shown in Table 15. The main effect of grade and sex both demonstrated significant differences (P < 0.001). The interaction between grade and sex was also significantly different for grade 2 (P<0.008), grade 6 (P<0.001) and grade 7 (P<0.031).

***Table 15:*** *20m Shuttle Run Test (stages) of children from Grades 1 to 7 (n=292). Values are expressed as mean ± SD.*

| Grade[a] | Female | N | Male | n | Total[b] | N |
|---|---|---|---|---|---|---|
| 1 | 3.9 ± 1.2 | 22 | 4.9 ± 1.4 | 28 | *4.5 ± 1.4* | *50* |
| 2 | 3.9 ± 1.4 | 18 | 5.7 ± 2.0 | 27 | *5.0 ± 2.0* | *45* |
| 3 | 4.2 ± 1.5 | 22 | 5.3 ± 2.1 | 24 | *4.8 ± 1.9* | 46 |
| 4 | 4.8 ± 1.7 | 15 | 5.9 ± 2.2 | 22 | *5.5 ± 2.1* | 37 |
| 5 | 5.2 ± 1.9 | 19 | 6.5 ± 2.6 | 23 | *5.9 ± 2.4* | 42 |
| 6 | 6.2 ± 1.9 | 14 | 7.9 ± 2.1 | 22 | *7.2 ± 2.2* | 36 |
| 7 | 5.8 ± 2.0 | 23 | 7.5 ± 2.4 | 13 | *6.4 ± 2.3* | 36 |
| *Total* | *4.8 ± 1.8* | *133* | *6.1 ± 2.3* | *159* | 5.5 ± 2.2 | *292* |

| | | |
|---|---|---|
| [a]*Main effect of Grade* | P < 0.001 ($1.1424^{-10}$) | $F_{6, 278} = 10.7$ |
| 1 vs 5, 6, 7 | *P < 0.005* | |
| 2 vs 6 | *P < 0.001* | |
| 3 vs 6, 7 | *P < 0.004* | |
| 4 vs 6 | *P < 0.002* | |
| [b]*Main effect of Sex* | P < 0.001 ($3.1825^{-9}$) | $F_{1, 278} = 37.4$ |
| Male vs Female | *P < 0.001* | |
| *Interaction Grade x Sex* | P < 0.896 | $F_{6, 278} = 0.4$ |

The TE and SWC arranged according to grade, and sex are shown in Table 16. The ability of the 20m shuttle run test to detect *small, medium* and *large* changes is summarised in this table. The SWC to detect *small* changes in the 20mSRT in each of the grades (and for both sexes) was less than the TE, showing a test rating of 'marginal'. The SWC to detect *medium* changes for Grade 6 females showed a testing rating of 'marginal'. Grade 1 and 2 males also showed a test rating of 'marginal'. The rest of the grades and sexes showed a SWC that had a 'good' ability to detect *medium* and *large* changes in the 20mSRT as the SWC was greater than the TE for each of these grades and sexes (Table 16).

***Table 16:*** *Ability of the 20m SRT to detect Small, Medium and Large Changes using Typical Error of Measurement (stages) and Smallest Worthwhile Change (stages) for 20m SRT (stages) of children from Grades 1 to 7 (n=289). Data expressed as mean (90% CI).*

| | | **TE** | **SWC** | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| **Sex** | **Grade** | | Small | Test Rating | Medium | Test Rating | Large | Test Rating | N |
| Females | 1 | 0.4 (0.4-0.5) | 0.3 | Marginal | 0.6 | Good | 1.0 | Good | 22 |
| | 2 | 0.6 (0.5-0.8) | 0.3 | Marginal | 0.7 | Good | 1.0 | Good | 18 |
| | 3 | 0.5 (0.5-0.7) | 0.3 | Marginal | 0.7 | Good | 1.0 | Good | 22 |
| | 4 | 0.7 (0.6-0.9) | 0.3 | Marginal | 0.7 | Good | 1.2 | Good | 15 |
| | 5 | 0.7 (0.6-0.9) | 0.3 | Marginal | 0.8 | Good | 1.4 | Good | 19 |
| | 6 | 1.1 (0.9-1.4) | 0.4 | Marginal | 0.9 | Marginal | 1.5 | Good | 14 |
| | 7 | 0.9 (0.8-1.1) | 0.4 | Marginal | 1.0 | Good | 1.6 | Good | 23 |
| | *Total* | *0.8 (0.7-0.9)* | *0.3* | *Marginal* | *0.8* | *Good* | *1.3* | *Good* | *133* |
| Males | 1 | 0.9 (0.8-1.1) | 0.3 | Marginal | 0.7 | Marginal | 1.2 | Good | 28 |
| | 2 | 1.0 (0.9-1.2) | 0.4 | Marginal | 0.9 | Marginal | 1.5 | Good | 27 |
| | 3 | 0.8 (0.7-1.0) | 0.4 | Marginal | 1.0 | Good | 1.6 | Good | 24 |
| | 4 | 0.7 (0.6-0.9) | 0.4 | Marginal | 1.1 | Good | 1.8 | Good | 22 |
| | 5 | 0.8 (0.6-0.9) | 0.5 | Marginal | 1.3 | Good | 2.1 | Good | 23 |
| | 6 | 1.0 (0.9-1.2) | 0.4 | Marginal | 1.1 | Good | 1.7 | Good | 22 |
| | 7 | 1.3 (1.0-1.7) | 0.5 | Marginal | 1.2 | Marginal | 1.9 | Good | 13 |
| | *Total* | *1.0 (0.9-1.0)* | *0.4* | *Marginal* | *1.1* | *Good* | *1.8* | *Good* | *159* |
| *All* | *1-7* | *0.9 (0.9-0.9)* | *0.4* | *Marginal* | *1.0* | *Good* | *1.7* | *Good* | *292* |

TE = Typical Error of Measurement; SWC = Smallest Worthwhile Change; Small = 0.2 × between subject SD; Medium = 0.5 × between subject SD; Large = 0.8 × between subject SD; TEM<SWC (Good); TEM=SWC (OK); TEM>SWC (Marginal); CI = Confidence Interval

Intraclass Correlation Coefficients (ICC) for each of the fitness test mesasurements are shown in the table below.

**Table 24:** Intraclass Correlation Coefficients for BMI, WC, SLJ, HGS and 20mSRT

| | ICC |
|---|---|
| BMI | 0.99 |
| WC | 0.96 |
| HGS | 0.95 |
| SLJ | 0.89 |
| 20mSRT | 0.82 |

The field-based tests that demonstrate strong/good predictive and criterion-related validity are the 20m shuttle run test, standing long jump, handgrip strength test, waist circumference and body mass index. Table 17 summarises these field-based tests' ability to detect *small, medium* and *large* changes.

***Table 17:*** *Summary of those field -based fitness tests that show strong / good predictive and criterion-related validity and their ability to detect small, medium and large measurement changes.*

| Health-Related Fitness Component | Strong Predictive Evidence and Good Criterion-Related Validity | Ability to Detect ***small***, ***medium*** and ***large*** changes |
|---|---|---|
| **Cardiorespiratory Fitness** | 20m Shuttle Run | Good ability to detect *medium* and *large* changes |
| **Musculoskeletal Fitness** | Hand Grip Strength | Good ability to detect *medium* and *large* changes |
| | Standing Long Jump | Good ability to detect *medium* and *large* changes |
| **Body Composition** | Body Mass Index | Good ability to detect *small, medium,* and *large* changes |
| | Waist Circumference | Good ability to detect *small, medium,* and *large* changes |

# CHAPTER 5: DISCUSSION

Understanding the “noise” of a measurement (TE) and SWC offers researchers and practitioners an additional level of analysis when determining the practical relevance of the observed changes (Buck and Lambert, 2022; Swinton et al., 2018) in physical fitness test measurements. As summarised in Table 17, we found that Hand Grip Strength, Standing Long Jump, 20m Shuttle Run, Body Mass Index, and Waist Circumference tests all exhibited strong predictive validity and good criterion-related validity.

When reviewing improvements or decrements in these physical fitness test results, it is important to identify what performance change is considered of practical significance (Moir et al., 2004; Spencer et al., 2006). The SWC is a good starting point to determine whether a practically significant change in performance has occurred (Hopkins, 2000; Pyne, 2003; Roe et al., 2016; Swinton et al., 2018). If the changes exceed the TE, we can be more confident that a real and practically significant change has occurred. Researchers and practitioners could utilise the SWC and TE as guidelines for setting targets when evaluating the degree of certainty concerning the practical significance of performance changes in a physical fitness characteristic (Buck and Lambert, 2022; Roe et al., 2016; Swinton et al., 2018).

TE and SWC can be used to identify real changes that represent practical significance in addition to statistical significance testing (Buck and Lambert, 2022; Hopkins, 2000; Moir et al., 2004; Roe et al., 2016; Spencer et al., 2006). Table 18 shows the various permutations that researchers may encounter when applying the TE and SWC approach to data from exercise training interventions or intervention studies (Buck and Lambert, 2022).

***Table 18:*** *Summary of the different permutations researchers may encounter when applying TE and SWC approaches to intervention studies.*

| | STATISTICALLY SIGNIFICANT | STATISTICALLY NOT SIGNIFICANT |
|---|---|---|
| PRACTICALLY SIGNIFICANT | 1 | 3 |
| NOT PRACTICALLY SIGNIFICANT | 2 | 4 |

1. Statistically significant ($P < 0.05$) and Practically significant (SWC > TE)
2. Statistically significant ($P < 0.05$) yet Not practically significant (SWC < TE)
3. Statistically not significant ($P \geq 0.05$) yet Practically significant (SWC > TE)
4. Statistically not significant ($P \geq 0.05$) and Not practically significant (SWC < TE)

It is worth discussing each of the physical fitness test results, along with the practical applications derived from our TE and SWC findings.

## Body Mass Index

The BMI results for males and females from Grades 1 to 7 are shown in Table 7. The mean and standard deviation for the combined group on the three testing days was 17.7 kg/m$^2$ (± 3.2 kg/m$^2$). The TE and SWC for each of the grades and sexes are shown in Table 9. The TE for the combined group is 0.2 kg/m$^2$. Since the *small (0.6* kg/m$^2$*)* SWC*, medium* (1.6 kg/m$^2$) SWC and *large* (2.5 kg/m$^2$) SWC are all larger than the TE, this demonstrates that the BMI test has a 'good' ability to detect *small, medium* and *large* changes.

Compared to the World Health Organisation growth reference data (De Onis et al., 2007), males in our study were 0.1% to 3% taller and 2.5% to 15% heavier than the reference data. Females were 0.3% to 4.3% taller and 13.4% to 21.8% heavier than the reference data. The body mass index was 0.3% to 10.3% higher in males and -1.1% to 13.7% higher in females compared to the reference data.

While numerous other studies have measured the BMI of male and female children, there does not appear to be any data on the TE and/or SWC of BMI in children aged 6 to 13 years. Therefore, a practical application of our TE and SWC data with other studies will follow. The permutations represented in Table 18 will be utilised to contextualise the results in a previous study by Di Maglie et al. (2022). Di Maglie et al.'s study investigated the effects of adding 40 minutes of physical activity 5-6 days per week for six months in a group of 12-year-old students. There were significant changes in BMI in the intervention group (-2.4 ± 0.6 kg/$m^2$) compared to the control group (+3.01 ± 1.8 kg/$m^2$).

***Table 19:*** *Summary of the SWC and TE data from the current study and how these values compare to the changes in BMI (kg/m2) from the Di Maglie et al. research study.*

| | **Our Research Data** | | | | **Di Maglie et al. Research Data** | |
|---|---|---|---|---|---|---|
| | SWC (small) | SWC (medium) | SWC (large) | TE | Exercise Group Changes | Control Group Changes |
| BMI (kg/$m^2$) | 0.5 | 1.2 | 1.9 | 0.1 | - 2.4 | + 3 |

If we apply the TE and SWC findings from our research (summarised in Table 21), the changes in the control and intervention groups would be considered *large* changes and are greater than the TE in 12-year-old (Grade 6) children. The changes in the intervention group were described as statistically significant by Di Maglie et al. and can be considered real changes of practical significance as they are greater than the TE of 0.1 kg/$m^2$.

## Waist Circumference

The mean and standard deviation of waist circumference was 62.9 cm (± 7.8 cm) for the combined group of males and females from Grades 1 to 7. The results for each of the 14 groups arranged according to sex and grade are summarised in Table 9. The TE for the combined group is 1.5 cm. All the grades and sexes showed a SWC with a 'good' ability to detect *medium* and *large* changes. However, the ability to detect *small* changes is 'marginal' in some grades across both sexes due to the TE being greater than the *small* SWC. The TE and SWC for each of the grades and sexes are shown in Table 10. A practical application using the TE and SWC data for waist circumference and the permutations represented in Table 18 applied to a previous study by Nickel et al. (2021) is described below. In the Nickel et al. study, the researchers investigated the effects of weekly lessons on healthy eating, physical activity and self-efficacy and two 30-minute structured physical education sessions in a group of Grade 1, 2 and 3 children. The authors reported significantly greater waist circumference reductions among the intervention participants than the control. A closer look at their data alongside our TE and SWC data for our combined group of grade 1, 2 and 3 children is shown in Table 22.

***Table 20:*** *Summary of the SWC and TE data from the current study and how these values compare to the changes that occurred in Waist Circumference (cm) from the Nickel et al. research study.*

| | **Our Research Data** | | | **Nickel et al. Research Data** | | |
|---|---|---|---|---|---|---|
| | SWC (small) | SWC (medium) | TE | Intervention Group Changes | Control Group Changes | Difference in Changes |
| WC (males) (cm) | 1.2 | 3.1 | 1.4 | - 0.7 | + 1.3 | 2.0 * |
| WC (females) (cm) | 1.4 | 3.4 | 1.4 | - 0.7 | + 0.6 | 1.3 ** |

* statistically significant change; ** no statistically significant change

On closer inspection of their data, the authors reported a statistically significant difference between the intervention and control groups of the male participants but no statistically significant difference between the female groups. If we apply our TE and SWC data, the males in their study would have shown a statistically significant and practically significant change (WC change > TE). In contrast, the changes observed in females can be described as neither statistically significant nor practically significant, as the TE was greater than the change. Consequently, this change could not be distinguished from the “noise” of the WC measurement.

## Hand Grip Strength (HGS)

*Small* changes in HGS for the combined group of grade 1-7 children were 1.3 kg. This was less than the TE of the measurement, showing the HGS test's ability to detect differences of this magnitude as “marginal.” However, the TE was less than the *medium* (3.3 kg) and *large* (5.3 kg) changes in HGS for all the grades and sexes, making changes of these magnitudes more interpretable as they exceeded the noise of the measurement. This information enables researchers to analyse the practical significance of the observed changes. The TE data can be described as homoscedastic, as the values did not increase as the HGS values increased. This makes it easier from a practical perspective since the TE can be expressed in the units of measurement (kg) rather than as a percentage (Table 12).

A previous study can demonstrate a practical application of the TE and SWC of the HGS test and the permutations in Table 18 (Siegel et al., 1989). This study assessed the effects of a 12-week group exercise program on prepubescent children (males and females) with a mean age of 8.4 ± 0.5 years. The authors concluded that a group strength training program can effectively increase fitness levels and improve body composition in males and females of this age. An examination of their data (refer to Table 19) showed that handgrip strength in the experimental group increased similarly on the right and left sides. However, despite a similar change, the authors found a statistically significant increase in the right handgrip but no statistically significant change in the left handgrip strength. A closer examination of their data revealed that the handgrip strength of the control group had increased more in the left hand of the males than in that of the right hand. The difference in the change between the control and experimental groups was, therefore, greater on the right than on the left side. This was the reason for the conclusion of a statistically significant change in right handgrip strength but no significant change in left handgrip strength.

***Table 21:*** *Summary of the SWC and TE data of 8-year-old males and females from our research and how these values compare to the changes in Hand Grip Strength from the Siegel et al. research study. Data expressed in kilograms (kg).*

| **HGS Measurement (kg)** | **Our Research Data (kg)** | | | **Siegel Research Data (kg)** | |
|---|---|---|---|---|---|
| | SWC (small) | SWC (medium) | TE | Experimental Group Changes | Control Group Changes |
| Right HGS (male) | 0.5 | 1.2 | 1.2 | + 1.5 | + 0.4 |
| Right HGS (female) | 0.6 | 1.4 | 1.2 | + 1.4 | + 0.4 |
| Left HGS (male) | 0.4 | 1.1 | 1.1 | + 1.2 | + 0.8 |
| Left HGS (female) | 0.6 | 1.5 | 1.3 | + 1.4 | + 0.0 |

Statistical significance does not always provide insight into whether a result is of practical significance, especially if a combination of small sample size and large measurement variability is apparent, as this can mask important effects (Batterham and Hopkins, 2006). The Grade 2 children in our study have a similar mean age, height and weight as the Siegel et al. study. The typical error (TE) or normal variation associated with the HGS test in 8-year-old males and females in our study is greater than the *small* SWC (Table 19), and therefore, one would not be able to detect such small changes as they cannot be perceived amongst the noise of the test. We can only be confident of a real and practically significant change if the data falls outside the calculated variation or typical error.

When we compare the improvements in handgrip strength from the Siegel et al. study with the SWC and TE of our study, the changes can be considered *medium* changes that are greater than the TE. The changes can be regarded as of practical relevance for both the right and left sides. Although the left-sided handgrip strength results were not statistically significant in the study by Siegel et al., they represent real changes of practical significance. Statistical significance does not provide information about the magnitude of the size effect and changes in handgrip strength greater than the TE and SWC can be considered real changes that are practically significant. Through this lens, researchers have an extra layer of analysis when determining the practical relevance of the observed changes.

## Standing Long Jump

Standing long jump changes associated with a *small* change for the combined group of Grades 1-7 children was 5.4 cm. This was less than the TE of the measurement (8.8 cm) showing the ability of the SLJ test as “marginal” for detecting differences of this magnitude. However, the *medium* (13.5 cm) and *large* (21.5 cm) changes in SLJ for all the grades and sexes were greater than the TE, making changes in these magnitudes more interpretable as they exceed the measurement's “noise.” The TE data is homoscedastic, as it did not increase as the SLJ values increased and is expressed in the units of measurement (cm) rather than as a percentage (Table 14).

Although other studies have measured the SLJ of male and female children (Chung et al., 2013; Thomas et al., 2020), there does seem to be no data on the TE and/or SWC of SLJ in children of ages 6 to 13 years. Fang et al. (2020) investigated the reliability, sensitivity and minimum detectable change (MDC) of SLJ in preschool children (ages 3.5 to 6 years). They showed a *small* SWC of 4.5 cm and concluded that the SLJ test had an excellent sensitivity to detect changes of this magnitude. Since the TE and SWC data from the current study cannot be compared directly to the study of Fang et al. (due to the differences in the population studied) or any other studies, a discussion of the practical applications of the findings is worth exploring in studies where SLJ changes are measured.

When reviewing improvements or decrements in SLJ test results, it is important to identify what change in performance is of practical significance (Moir et al., 2004; Spencer et al., 2006). The SWC is a good starting point to determine whether a practically significant change in performance has occurred (Buck and Lambert, 2022; Hopkins, 2000; Pyne, 2003; Roe et al., 2016; Swinton et al., 2018). The study results show *small* changes would be difficult to detect amongst the noise/variation (TE) of measurement. However, *medium* changes are larger than the TE and can be detected above the "noise" of the measurement. Therefore, if the changes exceed the TE, we can be more confident that a real and practically significant change has occurred. Researchers and/or practitioners could utilise the SWC and TE as guidelines to set targets when trying to ascertain the degree of certainty to which the SLJ performance change is practically significant (Buck and Lambert, 2022; Roe et al., 2016; Swinton et al., 2018).

The various permutations of Table 18 can once again be applied to the data collected from exercise training interventions or intervention studies (Buck and Lambert, 2022). A practical application of these guidelines can be demonstrated in a previous study (Faigenbaum et al., 2002). This study compared a group of children aged 7 to 12-years who participated in strength training once a week with a similar group that trained twice a week and a control group. They measured the effect of training on upper body strength, lower body strength and motor performance ability. Both training groups had significant gains in strength measured by a 1RM chest press and 1RM leg press. However, no statistically significant training-induced changes in standing long jump performance were observed.

Their data showed that standing long jump performance improved in the control group, one-day per week and two-day per week strength training group, by 7.5 cm, 7 cm and 10 cm, respectively. These changes in SLJ were not statistically significant. If we apply the TE and SWC findings from our research, the changes in the control and one-day per week strength training groups cannot be observed amongst the "noise" of the test. These changes could be described as statistically not significant and not practically significant. We can only be confident of a real and practically significant change if the data falls outside of the calculated variation or typical error. The two-day per week strength training group improved 10 cm, which is higher than our TE of 8.8cm in a group of children with similar mean age, height and weight. We could argue that despite a nonsignificant SLJ result in this group, the change of 10 cm is a real change of practical significance (Buck and Lambert, 2022).

Changes greater than the TE can be considered real, practically significant changes (Buchheit et al., 2011; Buck and Lambert, 2022; Hopkins, 2004; Hopkins, 2000; Pyne, 2003) in SLJ performance. Through this lens, these changes in the two-day per week strength training group could be considered practically significant but statistically not significant. Therefore, practitioners and researchers have an extra layer of analysis when determining the practical relevance of the observed changes in SLJ performance.

## 20m Shuttle Run Test (20mSRT)

Changes associated with a *small* change were 0.4 stages of a 20mSRT stage for the Grades 1-7 children group, which was less than the TE of measurement (0.9 stages). Thus, the 20mSRT test's ability to detect changes of this magnitude is "marginal". *Medium* (1.0 stages) and *large* (1.7 stages) changes in 20mSRT were greater than the TE; therefore, the test can detect changes of this magnitude for the group of grade 1-7 children. Many other studies have performed the 20mSRT with children. The current study showed similar results to those of Tottori et al. (2019), who performed in Japan, where children obtained a mean of 42.3 total shuttles in the 20mSRT. The current study demonstrated a mean of 44.5 total shuttles (converted from the mean of 5.5 stages). However, in contrast, the results of a study by van Stryp et al. (2022) performed with South African children showed a mean of 3.8 total stages amongst Grade 1 boys and girls, whereas our study showed a mean of 4.5 total stages. It is interesting to note that the mean height and body mass of the Grade 1's in the current study differed from that of the van Stryp et al.'s study with the height, body mass and BMI being significantly higher in the current study. Stryp et al. compared their data to normative values produced by Tomkinson et al. (2017) from 50 countries and concluded that the children in their study showed lower than average fitness levels compared to these normative values. The results of the current study show 20mSRT scores that are just above the 80$^{th}$ percentile compared to the normative data produced by Tomkinson's research, demonstrating significantly higher 20mSRT scores and fitness levels.

Other studies have measured the 20mSRT of male and female children however, there appears to be a lack of data on the TE and/or SWC of 20mSRT in children ages 6 to 13 years. We, therefore, cannot compare our TE and SWC data with other studies, so a discussion of the practical application of our TE and SWC data will follow. An application of the permutations represented in Table 18 will be demonstrated in a previous study by Tottori et al. (2019). In Tottori et al.'s study, the researchers investigated the effects of a high-intensity interval training (HIIT) program on physical fitness and executive function in a group of 8- to 12-year-old children. The children were divided into a HIIT group and a control group, with the 20mSRT being one of the tests used to assess changes in physical fitness levels.

Upon examining their data, the HIIT group demonstrated a statistically significant improvement of 3.15 shuttles following the intervention. To compare the change to our TE and SWC data, this change can be converted to a value of 0.4 stages in the 20mSRT (Table 20). This represents a '*small*' change, which would be difficult to see amongst the "noise" of the test as it is less than the TE of 0.9 stages shown in the current study in a group of children with similar mean age, height and body mass. Despite the researchers in the Tottori et al. study showing a statistically significant change in 20mSRT scores in the intervention group, this change is less than the TE and, therefore, represents a change that is not practically significant. This demonstrates once again that statistical significance does not always give adequate insight into whether a change is of practical relevance, especially as small sample size and large measurement variability can mask important effects (Batterham and Hopkins, 2006). TE and SWC allow researchers and practitioners to apply an additional layer of analysis when trying to determine the practical relevance of the observed changes in physical fitness tests.

***Table 22:*** *Summary of the SWC and TE data from the current study and how these values compare to the changes that occurred in 20mSRT (stages) from the Tottori et al. research study. Data expressed in stages.*

| | **Our Research Data** | | | **Tottori et al. Research Data** | |
|---|---|---|---|---|---|
| | SWC (*small*) | SWC (*medium*) | TE | HIIT Group Changes | Control Group Changes |
| **20m SRT (stages)** | 0.4 | 1.0 | 0.9 | + 0.4 | - 0.15 |

## Limitations and recommendations for further study

This study has limitations that must be taken into account when interpreting and applying the results. The school attended by the children in this study is classified as a Quintile 5 school in the South African education system (Grant, 2013). The South African Department of Basic Education defines the quintiles as follows: Quintile 1 comprises schools in each province (region/state) catering to the poorest 20% of learners. Quintile 2 schools serve the next poorest 20% of learners, and so forth. Quintile 5 schools cater to the least poor 20% of learners. Consequently, a Quintile 5 school would have learners from households with a higher socioeconomic status compared to the other quintiles. Therefore, caution must be exercised when generalising the results from our study to populations that differ significantly from the one examined.
A Jamar dynamometer was used to obtain our handgrip strength values, as this has been identified as the gold standard for measuring HGS (Gąsior et al., 2018). Gasior et al. (2018) found that the second handle position of the Jamar hand dynamometer is optimal for obtaining maximal HGS in non-athlete paediatric participants with normal development (Häger-Ross and Rösblad, 2002). Amaral et al. (2012) reported that using different grip dynamometers generates different grip strength values. This could be a limitation when applying our TE and SWC results for HGS to intervention studies utilising dynamometers.

Measurement error has two components of variability: systematic bias and random error. Atkinson et al. (1998) reported that the random error component is typically larger than that arising from systematic bias and that researchers can do relatively little to minimise random error, particularly when it is entirely attributable to inherent instrument variation. Systematic bias is a trend for repeated test measurements to differ in a particular direction (positive or negative). It can be affected by a learning effect, recovery time between tests, training effects and the level of motivation given. We could control for systematic bias by adhering to a strict testing protocol during our study. This included explanations and familiarisation protocols for the children, sufficient rest between repeated attempts, motivation level kept constant and repeated tests performed within 7 days and at similar times of day to limit the chance of training effects and the influence of circadian rhythm on the results. Since we could control for systematic bias in our study and random error is due to inherent mechanical (instrument) variation, it would be interesting to see whether the TE and SWC values differ when different measuring equipment (HGS dynamometers, body mass scales) is used. This is an area of research that could be considered in the future.

Another area of future research could be examining the feasibility of performing a battery of the recommended tests from the current research (HGS, SLJ, 20mSRT, BMI, WC) in a school setting. It would be interesting to establish the typical test administration time per child or group of children, the number of test administrators required, and the testing equipment resources required in a school setting.

# CHAPTER 6: CONCLUSION

This study aimed to measure the reliability of health-related and skill-related physical fitness tests for children aged 6 – 13 years to determine which tests are appropriate for inclusion in a programme for use in a primary school setting. Those field-based physical fitness tests that demonstrated strong predictive validity and good criterion validity from the literature and low measurement error from the reliability data are:

- Hand grip strength
- Standing long jump
- Body mass index
- Waist circumference
- 20m shuttle run

These fitness tests could be included in a battery of tests for implementation within the school setting. Researchers and practitioners can utilise the SWC and TE as guidelines to set targets when trying to ascertain the degree of certainty to which performance changes in these fitness tests are practically significant. Thresholds for the magnitude of changes (“small”, “medium", and “large”) will assist practitioners in identifying the smallest change that exceeds the “noise” (TE). TE and SWC can identify real changes that represent practical significance in addition to statistical significance testing (Buck and Lambert, 2022; Hopkins, 2000; Moir et al., 2004; Roe et al., 2016; Spencer et al., 2006; Swinton et al., 2018). This is useful when observing and reporting on the observed changes.

Detecting small changes earlier has several important practical benefits. For example, an earlier reaction can refine an intervention and make it more precise. Practitioners can also provide more regular feedback on progress using these “smaller to see” changes. By refining interpretation and being able to observe the smaller changes, we reduce the risk of making type 2 errors and missing differences that do exist. The SWC and TE from the current data can provide additional insight when analysing fitness test changes following intervention studies or activities. This allows researchers and practitioners to gain more insight into the practical relevance of observed changes.

## REFERENCES:

Amaral JF, Mancini M, Novo Júnior JM. Comparison of three hand dynamometers in relation to the accuracy and precision of the measurements. Rev Bras Fisioter. 2012 Jun;16(3):216-24.

Armstrong ME, Lambert MI, Sharwood KA, Lambert EV. Obesity and overweight in South African primary school children -- the Health of the Nation Study. S Afr Med J. 2006 May;96(5):439-44.

Atkinson G, Nevill AM. Statistical methods for assessing measurement error (reliability) in variables relevant to sports medicine. Sports Med. 1998 Oct;26(4):217-38.

Baker JL, Olsen LW, Sørensen TI. Childhood body-mass index and the risk of coronary heart disease in adulthood. N Engl J Med. 2007 Dec 6;357(23):2329-37.

Batterham AM, Hopkins WG. Making meaningful inferences about magnitudes. Int J Sports Physiol Perform. 2006 Mar;1(1):50-7.

Batty GD, Clark H, Morton SM, MacIntyre S, Leon DA. Intelligence in childhood and mortality, migration, questionnaire response rate, and self-reported morbidity and risk factor levels in adulthood-preliminary findings from the Aberdeen 'Children of the 1950s' study. J Epidemiol Community Health. 2002 Sep;56(suppl II):A1.

Buchheit M, Lefebvre B, Laursen PB, Ahmaidi S. Reliability, usefulness, and validity of the 30-15 Intermittent Ice Test in young elite ice hockey players. J Strength Cond Res. 2011 May;25(5):1457-64.

Buck R, Lambert MI. Getting a Grip on Strength Measurement in Children (6-13 Y): Impact of Typical Error of Measurement. Pediatr Exerc Sci. 2022;34:141-147

Byrd-Williams CE, Shaibi GQ, Sun P, Lane CJ, Ventura EE, Davis JN, Kelly LA, Goran MI. Cardiorespiratory fitness predicts changes in adiposity in overweight Hispanic boys. Obesity (Silver Spring). 2008 May;16(5):1072-7.

Castro-Piñero J, Artero EG, España-Romero V, Ortega FB, Sjöström M, Suni J, Ruiz JR. Criterion-related validity of field-based fitness tests in youth: a systematic review. Br J Sports Med. 2010 Oct;44(13):934-943.

Castro-Piñero J, Ortega FB, Artero EG, Girela-Rejón MJ, Mora J, Sjöström M, Ruiz JR. Assessing muscular strength in youth: usefulness of standing long jump as a general index of muscular fitness. J Strength Cond Res. 2010 Jul;24(7):1810-7.

Chatrath R, Shenoy R, Serratto M, Thoele DG. Physical fitness of urban American children. Pediatr Cardiol. 2002 Nov-Dec;23(6):608-12.

Chung L, Chow L, Chung J. Normative reference of standing long jump indicates sex difference in lower muscular strength of pubertal growth. Health. 2013(5): 6-11.

Cruz-León C, Expósito-Carrillo P, Sánchez-Parente S, Jimenez-Iglesias J, Borges-Cosic M, Cuenca-Garcia M, Castro-Pinero J. Feasibility and Safety of Field-Based Physical Fitness Tests: A Systematic Review. Sports Med Open. 2025 11, 8.

Cvejić D, Tamara P, Sergej MO. Assessment of physical fitness in children and adolescents. FU Phys Ed Sport. 2013 Nov; 135-145.

De Onis M, Onyango AW, Borghi E, Siyam A, Nishida C, Siekmann J. Development of a WHO growth reference for school-aged children and adolescents. Bull World Health Organ. 2007 Sep;85(9):660-7.

Di Maglie A, Marsigliante S, My G, Colazzo S, Muscella A. Effects of a physical activity intervention on schoolchildren fitness. Physiol Rep. 2022, 10:e15115.

Docherty D, Canadian Society for Exercise Physiology. Measurement in Pediatric Exercise Science. Champaign, IL: Human Kinetics; 1996.

España-Romero V, Artero EG, Jimenez-Pavón D, Cuenca-Garcia M, Ortega FB, Castro-Piñero J, Sjöstrom M, Castillo-Garzon MJ, Ruiz JR. Assessing health-related fitness tests in the school setting: reliability, feasibility and safety; the ALPHA Study. Int J Sports Med. 2010 Jul;31(7):490-7.

Faigenbaum AD, Milliken LA, Loud RL, Burak BT, Doherty CL, Westcott WL. Comparison of 1 and 2 days per week of strength training in children. Res Q Exerc Sport. 2002 Dec;73(4):416-24.

Fang H, Ho IMK. Intraday reliability, sensitivity, and minimum detectable change of national physical fitness measurement for preschool children in China. PLoS One. 2020 Nov 20;15(11):e0242369.

Fraser BJ, Blizzard L, Buscot MJ, Schmidt MD, Dwyer T, Venn AJ, Magnussen CG. The Association Between Grip Strength Measured in Childhood, Young- and Mid-adulthood and Prediabetes or Type 2 Diabetes in Mid-adulthood. Sports Med. 2021 Jan;51(1):175-183.

Freedman DS, Wang J, Maynard LM, et al. Relation of BMI to fat and fat-free mass among children and adolescents. Int J Obes (Lond). 2005;29:1-8.

Gabbett TJ, Nassis GP, Oetter E, Pretorius J, Johnston N, Medina D. The athlete monitoring cycle: a practical guide to interpreting and applying training monitoring data. Br J Sports Med. 2017;51(20):1451–1452.

García-Hermoso A, Cavero-Redondo I, Ramírez-Vélez R, Ruiz JR, Ortega FB, Lee DC, Martínez-Vizcaíno V. Muscular Strength as a Predictor of All-Cause Mortality in an Apparently Healthy Population: A Systematic Review and Meta-Analysis of Data From Approximately 2 Million Men and Women. Arch Phys Med Rehabil. 2018 Oct;99(10):2100-2113.e5.

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 Jul;49(7):1079-1094.

García-Hermoso A, Ramírez-Vélez R, García-Alonso Y, Alonso-Martínez AM, Izquierdo M. Association of Cardiorespiratory Fitness Levels During Youth With Health Risk Later in Life: A Systematic Review and Meta-analysis. JAMA Pediatr. 2020 Oct 1;174(10):952-960.

Garnett SP, Baur LA, Srinivasan S, Lee JW, Cowell CT. Body mass index and waist circumference in midchildhood and adverse cardiovascular disease risk clustering in adolescence. Am J Clin Nutr. 2007 Sep;86(3):549-55.

Gąsior JS, Pawłowski M, Williams CA, Dąbrowski MJ, Rameckers EA. Assessment of Maximal Isometric Hand Grip Strength in School-aged Children. Open Med (Wars). 2018 Mar 15;13:22-28.

Grant D. Media Release, WCED. 2013 Oct. Available at: https://wcedonline.westerncape.gov.za/comms/press/2013/74_14oct.html (Accessed: 02 November 2024).

Häger-Ross C, Rösblad B. Norms for grip strength in children aged 4-16 years. Acta Paediatr. 2002;91(6):617-25.

HAKSA. Healthy active kids South Africa 2018 report card. 2018.

HAKSA. Healthy active kids South Africa 2022 report card. 2022.

Hasselstrøm H, Hansen SE, Froberg K, Andersen LB. Physical fitness and physical activity during adolescence as predictors of cardiovascular disease risk in young adulthood. Danish Youth and Sports Study. An eight-year follow-up study. Int J Sports Med. 2002 May;23 Suppl 1:S27-31.

Henriksson H, Henriksson P, Tynelius P, Ekstedt M, Berglind D, Labayen I, Ruiz JR, Lavie CJ, Ortega FB. Cardiorespiratory fitness, muscular strength, and obesity in adolescence and later chronic disability due to cardiovascular disease: a cohort study of 1 million men. Eur Heart J. 2020 Apr 14;41(15):1503-1510.

Hestbaek L, Leboeuf-Yde C, Kyvik KO. Are lifestyle-factors in adolescence predictors for adult low back pain? A cross-sectional and prospective study of young twins. BMC Musculoskelet Disord. 2006 Mar 15;7:27.

Hopkins WG. How to interpret changes in an athletic performance test. *Sportscience* 8(58): 1–7,2004.

Hopkins WG. Measures of reliability in sports medicine and science. Sports Med. 2000 Jul;30(1):1-15.

Janz KF, Dawson JD, Mahoney LT. Increases in physical fitness during childhood improve cardiovascular health during adolescence: the Muscatine Study. Int J Sports Med. 2002 May;23 Suppl 1:S15-21.

Johnson MS, Figueroa-Colon R, Herd SL, Fields DA, Sun M, Hunter GR, Goran MI. Aerobic fitness, not energy expenditure, influences subsequent increase in adiposity in black and white children. Pediatrics. 2000 Oct;106(4):E50.

Klein DH, Mohamoud I, Olanisa OO, Parab P, Chaudhary P, Mukhtar S, Moradi A, Kodali A, Okoye C, Arcia Franchini AP. Impact of School-Based Interventions on Pediatric Obesity: A Systematic Review. Cureus. 2023 Aug 8;15(8):e43153.

Kolimechkov S. Physical fitness assessment in children and adolescents: A systematic review. Eur J Phys Educ. 2017 Apr 8.

Koutedakis Y, Bouziotas C, Flouris AD, Nelson PN. Longitudinal modeling of adiposity in periadolescent Greek schoolchildren. Med Sci Sports Exerc. 2005 Dec;37(12):2070-4.

Lang JJ, Belanger K, Poitras V, Janssen I, Tomkinson GR, Tremblay MS. Systematic review of the relationship between 20m shuttle run performance and health indicators among children and youth. J Sci Med Sport. 2018 Apr;21(4):383-397.

Marques A, Henriques-Neto D, Peralta M, Martins J, Gomes F, Popovic S, Masanovic B, Demetriou Y, Schlund A, Ihle A. Field-Based Health-Related Physical Fitness Tests in Children and Adolescents: A Systematic Review. Front Pediatr. 2021 Mar 5;9:640028.

Martinez-Millana A, Hulst JM, Boon M, Witters P, Fernandez-Llatas C, Asseiceira I, Calvo-Lerma J, Basagoiti I, Traver V, De Boeck K, Ribes-Koninckx C. Optimisation of children z-score calculation based on new statistical techniques. PLoS One. 2018 Dec 20;13(12):e0208362.

Mikkelsson LO, Nupponen H, Kaprio J, Kautiainen H, Mikkelsson M, Kujala UM. Adolescent flexibility, endurance strength, and physical activity as predictors of adult tension neck, low back pain, and knee injury: a 25 year follow up study. Br J Sports Med. 2006 Feb;40(2):107-13.

Minatto G, Barbosa Filho VC, Berria J, Petroski EL. School-Based Interventions to Improve Cardiorespiratory Fitness in Adolescents: Systematic Review with Meta-analysis. Sports Med. 2016 Sep;46(9):1273-92.

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 Nov;48(11):2577-2605.

Moir G, Button C, Glaister M, Stone MH. Influence of familiarization on the reliability of vertical jump and acceleration sprinting performance in physically active men. J Strength Cond Res. 2004 May;18(2):276-80.

Monasor-Ortolá D, Quesada-Rico JA, Nso-Roca AP, Rizo-Baeza M, Cortés-Castell E, Martínez-Segura A, Sánchez-Ferrer F. Degree of Accuracy of the BMI Z-Score to Determine Excess Fat Mass Using DXA in Children and Adolescents. Int J Environ Res Public Health. 2021 Nov 18;18(22):12114.

Moreno LA, Blay MG, Rodriguez G, et al. Screening performances of the International Obesity Task Force body mass index cut-off values in adolescents. J Am Coll Nutr. 2006;25:403-408.

Nickel NC, Doupe M, Enns JE, et al. Differential effects of a school-based obesity prevention program: A cluster randomized trial. Matern Child Nutr. 2021; 17:e13009.

Nyawose ZZ, Naidoo R, Christie C, et al. Results From South Africa's 2022 Healthy Active Kids' Report Card on Physical Activity, Body Composition Proxies, and Nutritional Status in Children and Adolescents. J Phys Act Health. 2024 Aug 8;21(9):861-871.

Ortega FB, Cadenas-Sánchez C, Sánchez-Delgado G, Mora-González J, Martínez-Téllez B, Artero EG, Castro-Piñero J, Labayen I, Chillón P, Löf M, Ruiz JR. Systematic review and proposal of a field-based physical fitness-test battery in preschool children: the PREFIT battery. Sports Med. 2015 Apr;45(4):533-55.

Ortega FB, Ruiz JR, Castillo MJ, Sjöström M. Physical fitness in childhood and adolescence: a powerful marker of health. Int J Obes (Lond). 2008 Jan;32(1):1-11.

Pietrobelli A, Faith MS, Allison DB, Gallagher D, Chiumello G, Heymsfield SB. Body mass index as a measure of adiposity among children and adolescents: a validation study. J Pediatr. 1998 Feb;132(2):204-10.

Plowman SA and Meredith MD. Fitnessgram/Activitygram Reference Guide (4th Edition). Dallas, TX: The Cooper Institute. 2013.

Pyne DB. Interpreting the results of fitness testing. Paper presented at: International Science and Football Symposium Melbourne, Australia, 2003.

Roe G, Darrall-Jones J, Till K, Phibbs P, Read D, Weakley J, Jones B. Between-days reliability and sensitivity of common fatigue measures in rugby players. Int J Sports Physiol Perform. 2016 11(5): 581–586.

Ruiz JR, Castro-Piñero J, Artero EG, Ortega FB, Sjöström M, Suni J, Castillo MJ. Predictive validity of health-related fitness in youth: a systematic review. Br J Sports Med. 2009 Dec;43(12):909-923.

Safrit MJ. Introduction to measurement in Physical Education and Exercise Science. St. Louis, MO: Ties Mirror/Mosby; 1986.

Safrit MJ, Costa MG, Hooper LM. The Validity Generalization Model: An Approach to the Analysis of Validity Studies in Physical Education. Res Q Exerc Sport. 1986 Aug; 57(4):288–297.

Semiz S, Ozgoren E, Sabir N. Comparison of ultrasonographic and anthropometric methods to assess body fat in childhood obesity. Int J Obes (Lond). 2007;31:53-58.

Siegel JA, Camaione DN, Manfredi TG. The Effects of Upper Body Resistance Training on Prepubescent Children. Pediatr Exerc Sci. 1989 May;1(2):145-154.

Silva D, Chaput J, Katzmarzyk P, et al. Physical education classes, physical activity, and sedentary behavior in children. Med Sci Sports Exerc. 2018;50(5):995–1004.

Simoni P. Body Composition. In: Simoni P, Aparisi Gómez MP, editor. Essential Measurements in Pediatric Musculoskeletal Imaging. Cham, Switzerland: Springer; 2023. p. 247-261.

Smith JJ, Eather N, Morgan PJ, Plotnikoff RC, Faigenbaum AD, Lubans DR. The health benefits of muscular fitness for children and adolescents: a systematic review and meta-analysis. Sports Med. 2014 Sep;44(9):1209-23.

Spencer M, Fitzsimons M, Dawson B, Bishop D, Goodman C. Reliability of a repeated-sprint test for field-hockey. J Sci Med Sport. 2006 May;9(1-2):181-4.

Swinton PA, Hemingway BS, Saunders B, Gualano B, Dolan E. A Statistical Framework to Interpret Individual Response to Intervention: Paving the Way for Personalized Nutrition and Exercise Prescription. Front Nutr. 2018 May 28;5:41.

Taylor RW, Jones IE, Williams SM, et al. Evaluation of waist circumference, waist-to-hip ratio, and the conicity index as screening tools for high trunk fat mass, as measured by dual-energy X-ray absorptiometry, in children aged 3-19y. Am J Clin Nutr. 2000;72:490-495.

Thomas E, Petrigna L, Tabacchi G, Teixeira E, Pajaujiene S, Sturm DJ, Sahin FN, Gómez-López M, Pausic J, Paoli A, Alesi M, Bianco A. Percentile values of the standing broad jump in children and adolescents aged 6-18 years old. Eur J Transl Myol. 2020 Jun 17;30(2):9050.

Tomkinson GR, Lang JJ, Blanchard J, Léger LA, Tremblay MS. The 20-m Shuttle Run: Assessment and Interpretation of Data in Relation to Youth Aerobic Fitness and Health. Pediatr Exerc Sci. 2019 May 1;31(2):152-163.

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(4):285-300.

Tottori N, Morita N, Ueta K, Fujita S. Effects of High Intensity Interval Training on Executive Function in Children Aged 8-12 Years. Int J Environ Res Public Health. 2019 Oct 26;16(21):4127.

Twisk JW, Kemper HC, van Mechelen W. The relationship between physical fitness and physical activity during adolescence and cardiovascular disease risk factors at adult age. The Amsterdam Growth and Health Longitudinal Study. Int J Sports Med. 2002;23(Suppl1):S8–14.

Van Stryp O, Africa E, Duncan M.J. Physical fitness and weight status of grade one children in Cape Town, South Africa. S Afr J Res Sport Phys Educ Recreation. 2022;44(2):81–90.

Washif JA, Pyne DB. Optimal Support for Elite Sprinters Using a Minimal, Adequate, and Accurate Approach. Int J Sports Physiol Perform. 2024 Sep 12:1-7.

Wind AE, Takken T, Helders PJ, Engelbert RH. Is grip strength a predictor for total muscle strength in healthy children, adolescents, and young adults? Eur J Pediatr. 2010 Mar;169(3):281-287.

# APPENDICES

## Appendix 1: Parent Information Sheet

**HEALTH / FITNESS ASSESSMENT OPPORTUNITY FOR YOUR CHILD (from Rob Buck, Sports Director)**

**What is this all about?**

I would like to invite your child to be involved in a series of fitness tests that will take place at Pinehurst in the first 2 weeks of the third term. The fitness tests will allow me to give you and your child feedback that can contribute to the improvement of their health and fitness.

**Why are you doing this?**

These fitness tests form part of a research study that I am performing during my Master's Degree in Exercise Science through the University of Cape Town. Myself and a research team are investigating which physical fitness tests are the most appropriate to use within the school setting.

**What do I need to do if I want my child to be involved?**

1). Read "Parent Information Sheet" (included in following pages)

2). Read and sign the "Parent Informed Consent" (included in following pages)

3). Complete the "Physical Activity Readiness Questionnaire" (included in following pages)

4). Read the "Child Information Sheet" with your child (included in following pages)

5). Let your child fill in their name on the "Child Assent Form" should they wish to take part

6). Return the following **completed** forms to school:

i). Parent Informed Consent

ii). Physical Activity Readiness Questionnaire

iii). Child Assent

7). Contact me on 084 981 0233 or research@robinbuck.co.za if you have any questions

**What do I need to do if I do NOT want my child to be involved?**

You do not need to do anything. Your child can only be part of the fitness tests and health / fitness feedback if you have given your consent and they agree to participate.

**PARENT INFORMATION SHEET**

Dear Parent / Guardian,

Our goal is to assess your child's health and physical development through a series of physical fitness tests. A research team from the University of Cape Town will be conducting a study which will contribute to this goal. The study will determine how best to conduct physical fitness tests within the school setting. These fitness tests have the potential to monitor and track your child's health and physical development. This is important as there is an established link between current fitness and future health in children. The study will allow us to give you and your child feedback that can contribute to the improvement of their health and fitness.

**Research Study title:** *A study to examine the validity and reliability of a selected group of physical fitness tests within a primary / elementary school setting*

The research team would like to invite your child to participate in this research study, which will be conducted during school time in the first two weeks of the third term. If you give permission, you will need to complete a Physical Activity Readiness Questionnaire (PARQ) on behalf of your child. This involves answering a few questions about your child's current health status that should take you about 5 minutes to complete. This form is included in this information pack.

The next part of the study involves your child performing specific physical fitness tests and measurements at the school. Measurements taken include weight, sitting and standing height and waist circumference. Leg and hand strength as well as flexibility and balance will also be measured. The fitness tests include the assessment of agility (ability to run as fast as possible and change direction), running speed and endurance capacity (ability to keep exercising for long periods).

The only discomfort your child may experience is during the endurance test where they have to run for as long as they can keep up. This may make them feel tired and short of breath. They may also experience some muscle stiffness a day or two after the fitness test, but this is a normal response, and the symptoms will disappear after a few days.

There is no financial compensation to you or your children for their participation. However, there will be the benefit of receiving a comprehensive printed physical fitness report that includes their results. This report will inform you of your child's current physical fitness levels and give an indication of whether these fall within healthy levels for their age.

The personal information and the physical fitness test results will remain confidential. Results from the assessments will be included in a scientific paper however no one will be able to identify you or your child from the results of the study. Upon request, the specific findings of the study will be provided.

If you decide to give permission for your child to participate in the study they will be volunteering to do so and will need to sign a form of assent. If you decide they cannot participate, you or your child will not be penalized in any way. You or your child can decide to stop participating in the study at any time during or after testing. Should you or your child choose to withdraw from the research study, there will be no penalty.

There are organizations that check that the research study was conducted in a humane and ethical manner, for example the Human Research Ethics Committee (HREC). Approval to conduct this study has been provided by the HREC and the Western Cape Department of Education.

**Who do I contact if I have questions about the study?**

Principal Investigator: Mr Rob Buck (research@robinbuck.co.za / 084 981 0233)

Supervisor: Professor Mike Lambert (mike.lambert@uct.ac.za / 021 650 4558)

The UCT's Faculty of Health Sciences Human Research Ethics Committee can be contacted at 021 406 6338 if you have any ethical concerns or questions about your child's rights or welfare as a participant in this research study.

Kind regards,

Rob Buck

## Appendix 2: Child Information Sheet

**CHILD INFORMATION SHEET**

Dear Participant,

I am completing a study with researchers from the University of Cape Town. We are trying to find out how best to use and explain the results from physical fitness tests performed at school.

These tests will help you become more aware of your health, physical and sporting development.

I would like to invite you to join this research study. If you agree, your parents will complete a Physical Activity Readiness Questionnaire (PARQ) for you. This involves answering a few questions about your health and should only take them about 5 minutes to complete.

In the next part of the study you will have to do some fitness tests and measurements at school. We will measure your weight, sitting and standing height and waist circumference. Then we will measure your leg and arm muscle strength as well as your flexibility and balance. We will also measure your agility *(ability to run as fast as possible and change direction)*, running speed and endurance *(ability to keep exercising for long periods).* The testing will be done during school hours in the last two weeks of the second term.

We expect you to manage the physical fitness tests because they are similar to performing sport or physical education (phys ed) activities. You may feel tired and short of breath during the endurance test, but as soon as you stop and recover these symptoms will go away. You may also have slight stiffness in your muscles, but this will disappear after a few days.

You will receive a report on your results, which will tell you which areas of your fitness may: “need some help” or are “nice and healthy” or are “really good” for your age and when compared with other children in the rest of the world.

If you decide to join the study you will be volunteering to do so and will not get paid. There are no problems if you decide you do not want to participate. You can also stop at any time during or after testing. None of the other children will see your results. No one will be able to identify you from the results of the study.

I hope you will join us for the research study. If you have any questions or queries, please come and have a chat to me at the sports office.

Kind regards,

Mr Buck

## Appendix 3: Parent Informed Consent Form

**PARENT INFORMED CONSENT**

**A study to examine the validity and reliability of a selected group of physical fitness tests within a primary / elementary school setting**

*A project conducted by Mr Robin Buck (Division of Exercise Science and Sports Medicine, Department of Human Biology, Faculty of Health Sciences, University of Cape Town)*

---

***SIGNATURE OF PARENT / GUARDIAN***

*I have read the information provided in the parent information sheet. I have been given an opportunity to ask questions and I hereby give permission for my child to take part in this study.*

Name of Parent: __________________ Name of Child: ____________________

Signature of Parent: _______________ Date: ______________________

---

The **UCT's Faculty of Health Sciences Human Research Ethics Committee** can be contacted on 021 406 6338 in case you have any ethical concerns or questions about your rights or welfare as a participant on this research study.

---

**'What happens if my child gets hurt while taking part in this study?'**

This research study is covered by an insurance policy taken out by the University of Cape Town if any participants suffer a bodily injury as a result of taking part in the study. The insurer will pay for all reasonable medical costs required to treat the bodily injury, according to the SA Good Clinical Practice Guidelines 2006.

***SIGNATURE OF INVESTIGATOR***

*Signature of Investigator:* ___________ *Date:* _____________________

## Appendix 4: Child Assent Form

**CHILD ASSENT**

**A study to examine the validity and reliability of a selected group of physical fitness tests within a primary / elementary school setting**

*A project conducted by Mr Robin Buck (Division of Exercise Science and Sports Medicine, Department of Human Biology, Faculty of Health Sciences, University of Cape Town)*

---

***CONSENT OF CHILD TO PARTICIPATE***

*I have read the information sheet about the study and have had an opportunity to ask questions. I agree to be a participant in this research study.*

*Child's Name (in own handwriting):* _____________________

*Date:* _______________

---

***SIGNATURE OF INVESTIGATOR***

*Signature of Investigator:* ____________ *Date:* ______________________

## Appendix 5: Physical Activity Readiness Questionnaire

**PHYSICAL ACTIVITY READINESS QUESTIONNAIRE**

**Completed by a Parent/Guardian of Child**

NAME OF CHILD: ………………………………………………………………………………

CHILD'S DATE OF BIRTH: ……………………………    CHILD'S AGE: ……………….

Your child will be performing physical fitness testing as part of the research study. Please complete the physical activity readiness questionnaire below. Should you have any queries related to any questions below, please contact the researcher via email (research@robinbuck.co.za), phone (084 981 0233) or at the information session.

| | ***Please tick appropriate box*** | |
|---|---|---|
| | **YES** | **NO** |
| Has the physical fitness testing that your child will participate in been fully explained to you? | ☐ | ☐ |

**Any information contained herein will be treated as confidential**

| | | |
|---|---|---|
| Has your doctor ever said that your child has a heart condition and that your child should only do physical activity recommended by a doctor? | ☐ | ☐ |
| Does your child ever experience chest pain during physical activity? | ☐ | ☐ |
| Does your child ever lose balance because of dizziness or do they ever lose consciousness? | ☐ | ☐ |
| Does your child have a bone or joint problem that could be made worse by a change in their physical activity participation? | ☐ | ☐ |

Does your child have uncontrolled asthma (i.e. asthma that is not easily controlled by an inhaler?) ☐ ☐

Is your doctor currently prescribing any medication for your child's blood pressure or a heart condition? ☐ ☐

Do you know of any other reasons why your child should not undergo physical activity? This might include diabetes, a recent injury, or serious illness. ☐ ☐

If you have answered **NO** to all questions then you can be reasonably sure that your child can take part in the physical activity requirement of this research study.

I ……………………………. declare that the above information is correct at the time of completing this questionnaire on date ……/……/…….

---

**Please note:**

Should your child's health status change so that you can answer "YES" to any of the questions on the previous page, please notify the investigators / researchers and consult with your doctor regarding the level of physical activity in which your child can participate.

---

**If you answered YES to one or more questions:**

Talk to your doctor in person discussing with him/her those questions to which you answered yes.

Ask your doctor if your child is able to participate in the physical fitness tests in the research study.

Doctor's Name…. ……………………………. Date …………………………..

Doctor's Signature …………………………….

---

Signature of Investigator / Researcher……………………….. Date …………………..

**Appendix 6: Ethics Letter**

UNIVERSITY OF CAPE TOWN

Faculty of Health Sciences

Human Research Ethics Committee

Room E53-46 Old Main Building

Groote Schuur Hospital

Observatory 7925 Telephone [021] 406 6492

Website: **Email:** sumayah.ariefdien@uct.ac.za
www.health.uct.ac.za/fhs/research/humanethics/forms

---

05 December 2017

HREC REF: 713/2017

Prof M Lamberts

Sports Science Institute Human Biology-

Boundry Road, Newlands

Dear Prof Lamberts

**PROJECT TITLE: A CROSS-SECTIONAL STUDY TO EXAMINE THE VALIDITY, RELIABILITY**

**AND FEASIBILITY OF PERFORMING THE SSISA SCHOOL FITNESS ASSESSMENT PROTOCOL TO ASSESS HEALTH-RELATED AND SKILL-RELATED PHYSICAL FITNESS WITHIN A SCHOOL SETTING (MSc Candidate - Mr R Buck)**

Thank you for your response letter dated 20 November 2017, addressing the Issues raised by the Sciences Human Research Ethics Committee (HREC).

It is a pleasure to inform you that the HREC has formally approved the above-mentioned study.

**Approval Is granted for one year until the 30 December 2018.**

Please submit a progress form, using the standardised Annual Report Form if the study continues beyond the approval period. Please submit a Standard Closure form if the study Is completed within the approval period.

(Forms can be found on our website: www.health.uct.ac.za/fhs/research/humanethlcs/forms)

**We acknowledge that the student: Robin Buck will also be Involved In this study.**

**Please quote the HREC REF in all your correspondence.**

Please note that the ongoing ethical conduct of the study remains the responsibility of the principal investigator.

Please note that for all studies approved by the HREC, the principal investigator **must** obtain appropriate Institutional approval, where necessary, before the research may occur.

Yours sincerely

PROFESSOR M BLQCKMAN

CHAIRPERSON, FHS HUMAN RESEARCH ETHICS

Federal Wide Assurance Number: FWA00001637.

Institutional Review Board (IRB) number: IRB00001938

This serves to confirm that the University of Cape Town Human Research Ethics Committee complies to the Ethics Standards for Clinical Research with a new drug in patients, based on the Medical Research Council (MRC-SA), Food and Drug Administration (FDA-USA), International Convention on

Harmonisation Good Clinical Practice (ICH GCP), South African Good Clinical Practice Guidelines (DoH 2006), based on the Association of the British Pharmaceutical Industry Guidelines (ABPI), and Declaration of Helsinki (2013) guidelines.

The Human Research Ethics Committee granting this approval is in compliance with the ICH Harmonised Tripartite Guidelines E6: Note for Guidance on Good Clinical Practice (CPMP/ICH/135/95) and FDA Code Federal Regulation Part 50, 56 and 312.