University of Groningen

Field-Based Tests of Strength and Anaerobic Capacity Used in Children With Developmental Coordination Disorder

Aertssen, Wendy; Jelsma, Dorothee; Smits-Engelsman, Bouwien

Published in:
Physical Therapy

DOI:
10.1093/ptj/pzaa118

IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

Document Version
Publisher's PDF, also known as Version of record

Publication date:
2020

Link to publication in University of Groningen/UMCG research database

Citation for published version (APA):
Aertssen, W., Jelsma, D., & Smits-Engelsman, B. (2020). Field-Based Tests of Strength and Anaerobic Capacity Used in Children With Developmental Coordination Disorder: A Systematic Review. *Physical Therapy*, *100*(10), 1825–1851. https://doi.org/10.1093/ptj/pzaa118

Copyright
Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

The publication may also be distributed here under the terms of Article 25fa of the Dutch Copyright Act, indicated by the “Taverne” license. More information can be found on the University of Groningen website: https://www.rug.nl/library/open-access/self-archiving-pure/taverne-amendment.

Take-down policy
If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from the University of Groningen/UMCG research database (Pure): http://www.rug.nl/research/portal. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.

Review

Field-Based Tests of Strength and Anaerobic Capacity Used in Children With Developmental Coordination Disorder: A Systematic Review

Wendy Aertssen, Dorothee Jelsma, Bouwien Smits-Engelsman

Objective. Children with developmental coordination disorder (DCD) are reported to have lower levels of strength and anaerobic capacity. The purpose of this study was to (1) identify field-based tests for strength and anaerobic capacity used in studies comparing children with DCD and children who were typically developing (TD), (2) examine the methodological quality of studies reporting psychometric properties and rate the psychometric properties of the examined test, and (3) summarize available evidence by combining the methodological quality of the studies and the quality of the psychometric properties of the test.

Methods. An electronic search was conducted in July 2019 in 4 electronic databases. For purpose 1, primary studies were included with no exclusion of study design in which children aged 4 to 18 years with DCD were compared with children who were TD on strength and/or anaerobic capacity measures. For purpose 2, primary studies were included with no exclusion of study design in which a psychometric property was investigated. The Consensus-Based Standards for Selection of Health Measurement Instruments (COSMIN) was used to evaluate the methodological quality of the 34 studies and rate the psychometric properties of the tests used.

Results. Hand-held dynamometer, bent knee push-up, vertical jump, standing long-jump, functional strength measurement, fitness test, and test battery can be recommended for TD, and the shuttle run item of the Bruininks-Oseretsky Test of Motor Proficiency-Second Edition and 10 m × 5 m sprints (straight and slalom) can be recommended for DCD.

Conclusion. Information regarding psychometric properties of field-based tests for strength and anaerobic capacity in children with DCD is lacking.

Impact. Information about the psychometric properties of field-based tests for strength and anaerobic capacity in children with DCD is lacking. More information is available on TD children, but it is also not complete; information regarding validity and responsiveness, especially, is missing. When using measures in children with DCD, it is important to keep in mind this lack of evidence for the validity and reliability of the outcomes for this target group.

W. Aertssen, PT, PhD, Avansplus, Physical Therapy Department, University of Professionals, Claudius Prinsenlaan 140, Breda, 4818 CP, The Netherlands. Address all correspondence to Dr Aertssen at: wendyaertssen@hotmail.com.

D. Jelsma, PT, PhD, KU Leuven, Developmental and Clinical Neuropsychology, University of Groningen, Groningen, The Netherlands.

B.C.M. Smits-Engelsman, PhD, Department of Health and Rehabilitation Sciences, Division of Physiotherapy, Faculty of Health Sciences, University of Cape Town, Rondebosch, Western Cape, South Africa.

[Aertssen W, Jelsma D, Smits-Engelsman B. Field-based tests of strength and anaerobic capacity used in children with developmental coordination disorder: a systematic review. *Phys Ther.* 2020;100:1825–1851.]

© The Author(s) 2020. Published by Oxford University Press on behalf of the American Physical Therapy Association. All rights reserved. For permissions, please email: journals.permissions@oup.com

Published Ahead of Print: July 6, 2020
Accepted: May 26, 2020
Submitted: August 29, 2019

Post a comment for this article at: *https://academic.oup.com/ptj*

Downloaded from https://academic.oup.com/ptj/article/100/10/1825/5868019 by University of Groningen user on 16 October 2020

Children need sufficient physical fitness to be able to participate fully in everyday sport activities and games.[1,2] Participation in these activities offers opportunities for having fun, interacting socially, and possibly promoting physical well-being.[3,4] Children's patterns of exertion during sport and play are characterized by short periods of intensive physical activity interspersed with short periods of reduced or less intensive activity.[5] These short bursts of physical activity, such as jumping, pushing a friend on a swing, running short distances when playing tag, or climbing a climbing frame, utilize muscle strength, power, and endurance and mostly engage the anaerobic energy system. Reduced levels of physical fitness appear to be a risk factor for an increase in weight and, over time, may lead to health problems in adolescence.[3,6] It is therefore important to identify children who experience problems in outdoor play and/or sport activities at an early age and analyze which underlying deficits are hindering them. Early identification may result in early intervention, which aims to improve participation in sport and playground activities and may prevent secondary health problems.

One group of children known to experience problems in performing activities of daily life, play, and sport activities are those with developmental coordination disorder (DCD).[7] DCD is a diagnosis for children who experience motor coordination problems during development. The level of motor skill is substantially below the level expected for the chronological age. Besides motor coordination problems, children with DCD are also known to have lower levels of physical fitness, including cardiorespiratory fitness, anaerobic capacity, and muscle strength.[2,8–17] Hence, international clinical practice recommendations state that not only should the level of motor performance be assessed but also the level of strength and anaerobic capacity.[7] Whether the measured level of anaerobic fitness performance actually reflects a lower level of anaerobic capacity in children with DCD is still unclear but can be expected. Several reasons can be suggested why children with DCD score lower on tests of anaerobic capacity. Avoiding sport and play activities will result in fewer training opportunities in daily life compared with their peers,[18,19] leading to lower anaerobic capacity. However, as mentioned in previous research regarding aerobic fitness performance in children with DCD, it does not necessarily have to be associated with lower capacity but may be due to psychological aspects such as the child's negative perception of their abilities or feeling of exertion.[18] Other explanations could lie in the fact that the movements used in different tests have high coordination or balance demands (throwing and jumping) or are less mastered by the child with DCD (running or cycling). It can be argued that tests for anaerobic capacity and strength measure a different construct in children with DCD than in well-coordinated children. So far, a few intervention studies aimed at only children with DCD presented a significant improvement in strength and/or anaerobic capacity measures.[20–24] However, if the change after intervention is similar to the measurement error of the test, the effect of intervention is still unclear. Hence, testing anaerobic capacity and muscle strength in children with DCD using psychometrically sound tools is important because both clinical and scientific conclusions are based on the results of such tests.

There are several ways of measuring strength and anaerobic capacity. Muscle strength can be defined as the maximal strength a muscle group can generate and is usually expressed in newtons or the 1-repetition maximum, which is based on the weight that can be lifted just once when performing a movement through full range of 1 or more body segments. Anaerobic capacity can be divided into muscle power, muscle endurance, and general anaerobic capacity. Muscle power is the product of force and velocity per unit of time. During a high-intensity, short-term power task—for example, throwing, jumping, or a short sprint—energy is released by breaking down high-energy molecules (adenosine triphosphate molecule and phosphocreatine molecule) and glucose (the lactic acid system) in the muscle cells without the use of oxygen. Muscle endurance can be defined as the ability to sustain or repeat a contraction as many times as possible in a certain limited timeframe.[25] Sprinting tests are commonly used to assess general anaerobic capacity.

Laboratory tests of (an)aerobic capacity and strength are recognized as gold standards because of strict standardization and because equipment like cycle ergometers enable collection of accurate and reliable data. Field-based tests are gradually gaining attention worldwide because they are easy to implement in large epidemiological studies. In clinical settings, therapists use field-based tests, like running or jumping tests, since these provide more specific information about outdoor play and sport. Field-based tests are portable, low cost, ecologically valid, and provide outcomes that often relate well to outcomes from most laboratory tests.[26] Given that outdoor play and sport activities require a considerable level of strength and anaerobic capacity, a critical evaluation of the psychometric properties of the existing strength and anaerobic field-based tests for children in general and in a specific target group like DCD is warranted.

The taxonomy of the Consensus-Based Standards for the Selection of Health Measurement Instruments (COSMIN) is a well-established tool to verify the psychometric properties of tests.[27] One of the psychometric properties is reliability. The reliability of a test is the extent to which it is free from measurement error and covers internal consistency, test-retest reliability, inter-and intra-rater reliability, and measurement error. The validity of a test implies the extent to which a test provides outcomes, which are supposed to measure according to the content, construct, and criterion validity of the test. A third important psychometric property is responsiveness; this

Downloaded from https://academic.oup.com/ptj/article/100/10/1825/5868019 by University of Groningen user on 16 October 2020

refers to the ability of a test to detect changes in performance over time.

Taken together, there is a need for an overview of the variety of field-based tests and their psychometric properties, which are used thus far to evaluate strength and anaerobic capacity of children with DCD compared with children with TD, so results in daily practice and in scientific research can be interpreted correctly. Therefore, the purpose of this study is to (1) identify field-based tests for strength and anaerobic capacity used in studies comparing children with DCD and TD, (2) examine the methodological quality of studies reporting psychometric properties and rate the psychometric properties of the examined test, and (3) summarize available evidence by combining the methodological quality of the studies and the quality of the psychometric properties of the test.

Methods

Data Sources and Search

A comprehensive electronic search for relevant studies was conducted in July 2019 from the following electronic databases: Cochrane Database of Systematic Reviews; Cochrane Central Register of Controlled Trials; the Physiotherapy Evidence Database; and US National Library of Medicine Database (PubMed). The reference lists of all included articles were manually searched for additional studies.

Part 1. In the first search, we selected for constructs of interest (strength and anaerobic capacity) and target population (children with DCD). The inclusion criteria were (1) primary studies published between 1980 and 2019 with no exclusion based on study design; (2) used field-based tests intending to measure strength and/or anaerobic capacity; (3) were published in peer-reviewed journals in the English language; (4) included children aged 4 to 18 years with the clinical diagnosis of DCD (meeting all the criteria for DCD according to the Diagnostic and Statistical Manual of Mental Disorders [DSM 4 or DSM 5][28] or with p-DCD [not formally diagnosed or not all the criteria of the DSM 4 or DSM 5 are described in the study]) or children with mild motor problems and were compared with children with TD.

Part 2. In the second part, we searched for the psychometric properties of the previously identified field-based tests (Part 1) for strength and anaerobic capacity in children with TD and/or children with DCD by entering names of all the identified tests and psychometric properties (eg, reliability, validity, and responsiveness). The inclusion criteria were studies that (1) were primary studies with no exclusion for type of study design, (2) were published in peer-reviewed journals in the English language, (3) included children aged 4 to 18 years, and (4) reported on at least 1 psychometric property.

Studies focusing on children with other medical or neurological conditions such as cerebral palsy, cystic fibrosis, and rheumatic conditions were excluded.

For the search terms, see the Appendix.

Selection Process

In conducting this review, 1 author undertook the database search (W.A.). First, titles and abstracts of articles were screened independently by 2 authors (W.A. and D.J.) to select the relevant studies meeting the inclusion criteria. Any disagreement between the 2 authors was discussed until consensus was reached. Next, full texts of potentially relevant studies were obtained and reviewed independently by 2 authors (W.A. and D.J.). Any disagreements in scoring between the 2 authors were resolved after consensus or discussion. If needed, a third author was asked for advice (B.S.E.).

Data Extraction

Data extraction was conducted independently by 2 authors (W.A. and D.J.). For the first part of our study, we extracted information as follows: publication details (first author, year of publication, title, country); study design; information regarding study methods (inclusion/exclusion criteria for participants, recruitment of participants); participant characteristics (number of participants, age); tests for strength and anaerobic capacity; intervention details, when present (name of intervention, aim of intervention, number of intervention/control groups, intervention components, duration of intervention, primary and secondary outcomes, time points when data were collected, intervention effects on primary and secondary outcomes); and relevant author comments, for example, sensitivity towards change and references for psychometric properties of the tests used. For the second part of the study, the following information was extracted from the included studies: publication details (first author, year of publication, title, country), information regarding study methods (inclusion/exclusion criteria for participants, recruitment of participants, study design), participant characteristics (number of participants, age, country), tests for strength and anaerobic capacity used, and statistics regarding psychometric properties.

Quality Assessment Psychometric Properties

The quality of psychometric properties is a 3-step procedure according to COSMIN.[27] COSMIN is commonly used in systematic reviews regarding measurement properties. The first step is to assess if the study uses the right methodology for a specific psychometric property and qualify the design as very good, adequate, doubtful, inadequate, or not applicable. The second step is to rate the psychometric properties of the specific test based on

Downloaded from https://academic.oup.com/ptj/article/100/10/1825/5868019 by University of Groningen user on 16 October 2020

that study. The third step is combining the methodological quality of the studies, including studies rated at least "adequate," with the rated psychometric qualities of the test based on the robustness of the evidence.

Methodological quality assessment of included studies. The selected studies were investigated independently by 2 reviewers (W.A. and D.J.) for methodological quality by using the risk of bias checklist of the COSMIN.[27] The COSMIN risk of bias checklist consists of 9 boxes (content validity, structural validity, internal consistency, cross-cultural validity, reliability, measurement error, criterion validity, construct validity, and responsiveness), each describing the methodological quality to study that specific measurement property. The scoring of each item in those boxes is conducted using a 4-point rating scale indicating if the methodology used for the item in that box was inadequate, doubtful, adequate, or very good. Per box, the "lowest" rating leads to the conclusion of that specific box.

Criteria for evaluating psychometric properties of included tests. In step 2, we rated the researched psychometric properties of a specific test (for instance the reliability) based on the results of a study using the criteria for measurement properties as described at the COSMIN website.[27] Scores could be positive (+) when meeting all the described criteria for that property, negative (−) when the criteria were not met, or indeterminate (?) when information was lacking or unclear (Tab. 1).[29]

Summarized Evidence

In Step 3, we summarized the available evidence by combining the methodological quality of the studies (step 1) and the quality of the psychometric properties of the test (step 2). In principle, an instrument with good psychometric properties could be studied using an inadequate design or inappropriate statistics. Therefore, only studies with adequate or good methodological quality were used in this summary.[27] By performing this 3-step procedure, it is possible to draw conclusions because psychometric evidence reported was gathered in studies using the right methodology for that specific property.

Results

Part 1

In the first search, we identified 1075 papers in the different databases and another 8 by manual search. Duplicates were removed and titles and abstracts reviewed by 2 reviewers to determine if the study contained field-based tests for strength and anaerobic capacity in children with DCD. After reading the full text of the remaining 66 papers, 43 papers were excluded because they did not meet the inclusion criteria (2 reviews, 6 different population, 8 did not compare DCD and TD, 1 was not a primary study, 26 studies did not use field-based tests or strength/anaerobic capacity tests). Finally, 23 studies were included (Suppl. Fig. 1). Four of the 23 studies were intervention studies with pre- and posttest measurements with children with DCD and children with TD.[30–33]

The results of the different studies are reported in Table 2. In 21 of the 23 studies, differences in performance between children with DCD and children with TD were found. The measures used were able to discriminate between these 2 groups of children, supporting the known group validity of the measures.

Part 2

In the second search, we identified 660 papers and 1 paper by manual search. After reviewing the titles and abstracts, 599 studies were excluded for not meeting the inclusion criteria. Two reviewers (W.A. and D.J.) read the full texts of the remaining 62 studies and excluded another 28 studies because they did not meet the inclusion criteria (8 reviews, 3 wrong age, 2 test manuals, 7 studies did not use field-based tests or strength/anaerobic capacity tests, 5 did not include psychometric properties, 3 studies were not available) (Suppl. Fig. 1).

Methodological Quality of Included Studies

The 34 included studies were assessed for methodological quality using the COSMIN risk of bias checklist. Initially, there was a 73% agreement on the methodological quality between the 2 reviewers (W.A. and D.J.). After discussion, 100% consensus was reached. Of all the COSMIN boxes scored (content validity, structural validity, internal consistency, cross-cultural validity, reliability, measurement error, criterion validity, construct validity, and responsiveness), 73% of the studies had adequate to very good quality and 27% scored as doubtful or inadequate. Of the 34 studies, only 18 studies could be scored on the box for measurement error, resulting in the score of 10 studies as very good, 3 studies as adequate, 1 study as doubtful, and 4 studies as inadequate. Table 3 shows the methodological qualities of the different study designs to examine the psychometric properties according to the risk of bias checklist of COSMIN.

Criteria for Evaluating Psychometric Properties of Included Tests

Tables 4 and 5 show the results of the psychometric properties per test. Initially, 78% agreement was reached between the 2 reviewers (W.A. and D.J.) on the criteria for good measurement properties. After discussion, 100% consensus was reached.

Population. In the 34 studies in part 2 of our review, 12,450 children with TD and 129 children with DCD were included (aged 4–18 years). Of these 34 studies, only 7 studies included children older than 12 years. The studies were performed in 11 different countries (China, France,

Downloaded from https://academic.oup.com/ptj/article/100/10/1825/5868019 by University of Groningen user on 16 October 2020

Table 1.
Quality Criteria for Measurement Properties[a] [29]

Psychometric Property	Positive (+)	Indeterminate (?)	Negative (−)
Reliability	ICC > 0.70 Kappa > 0.70	ICC or Kappa not reported	ICC < 0.70 Kappa < 0.70
Measurement error	SDC or LoA < MIC	MIC not defined	Criteria + not met
Content validity	All items refer to relevant aspects of construct to be measured AND are relevant for target population AND for purpose of measurement instrument AND together comprehensively reflect construct to be measured	Not all information for + provided	Criteria + not met
Structural validity	Unidimensionality: First factor >20% of variability AND ratio of variance explained by first to second factor >4	Not all information for + provided	Criteria + not met
Internal consistency	Positive structural validity AND Cronbach's alpha >.7 and <.95	Not all information for + provided OR conflicting or negative evidence for structural validity	Criteria + not met
Criterion validity	Rationale for golden standard is clear AND correlation with gold standard ≥0.70	Not all information for + provided	Criteria + not met
Construct validity	At least 75% of results accord with hypotheses	No correlation with instrument measuring same construct AND no differences between relevant groups reported	Criteria + not met
Responsiveness	At least 75% of results accord with hypotheses	No correlations with changes in instrument measuring related construct AND no differences between changes in relevant groups reported	Criteria + not met

[a] ICC = interclass correlation; LoA = limits of agreement; MIC = minimal important change; SDC = smallest detectable change.

the Netherlands, Colombia, United States, Spain, Canada, Greece, Norway, South Africa, Australia) on 5 different continents.

Strength and anaerobic capacity tests. The different tests were divided into isometric strength, muscle power, muscle endurance, sprint test, and test batteries. The psychometric properties are reported in Tables 4 and 5. Structural validity was reported in only 2 studies. It was investigated for the Functional Strength Measurement (FSM; 64% explained by 1 factor) with positive results[46] and for the Bruininks Oseretsky Test-Short Form (BOT-SF)[51] with negative results.

Table 6 provides the summarized evidence of the different measures based on methodological quality (only studies scored adequate or very good were included) and the criteria of good measurement properties.

Discussion

The main aim of this systematic review was to (1) identify field-based tests for strength and anaerobic capacity used in studies comparing children with DCD and children with TD, (2) examine the methodological quality of studies reporting psychometric properties and rate the psychometric properties of the examined test, and (3) summarize the available evidence by combining the methodological quality of the studies and the quality of the psychometric properties of the test.

First, it was concluded that, although there have been many studies investigating physical fitness in children with DCD, information about the psychometric properties of the tests used to assess children with DCD is lacking. The importance of the psychometric properties in a specific target group is highlighted in the only study of children with DCD.[50] In this study, differences in the psychometric properties were found between a TD and a DCD group. Pathophysiological constraints in children with DCD may have influenced test performance.[11] The difficulties children with DCD experience with balance, coordination, and fast repetitive movements may influence the reliability and especially the validity of a measure. One must consider that movements with a change in direction are more challenging for children with DCD compared with children with TD. Therefore, evidence-based recommendations for strength and anaerobic capacity measures cannot be made for children with DCD.

Secondly, it was concluded that psychometric properties of strength and anaerobic capacity field-based tests have

Downloaded from https://academic.oup.com/ptj/article/100/10/1825/5868019 by University of Groningen user on 16 October 2020

Table 2.
Studies Examining Strength and Anaerobic Capacity in Children With DCD and Children With TD[a]

Study	Design	Population No. (Age)	Outcome Measures Regarding Strength and Anaerobic Capacity	Intervention	Results
Aertssen et al, 2016a[8]	Cross sectional/comparative	48 DCD, 110 TD (7–10 y)	FSM, MPST		Children with DCD had significantly poorer performances on MPST and on items of FSM (except overarm and underarm throwing and chest pass) and a group × age effect for items of lower extremities, muscle endurance, and total score
Beutum et al, 2013[34]	Comparative	9 DCD, 9 TD (7–11 y)	BOT-2 (subtest strength)		Children with DCD performed significantly worse on subtest strength of BOT-2
Bonney et al, 2017b[30]	Pre-post single blinded design	57 DCD, 54 TD (6–10 y)	5 items of FSM (long jump, lateral step-up, sit-to-stand, stair climbing, and lifting a box), BOT-2 (subtest running speed and agility), 10 × 5 m sprint and 10 × 5 m slalom	20 min playing 2×/wk for 5 wk on Wii. One group with only ski slalom, other group with variable games	Children with DCD performed significantly worse on different physical fitness outcome measures. After intervention, DCD and TD improved significantly (except 10 × 5 m sprint in DCD group) on different measures.
Cairney et al, 2015[35]	Study protocol	300 p-DCD, 300 TD (4–5 y)	Standing long jump		
Cermak et al, 2015[36]	Cross-sectional comparative	53 DCD, 65 TD (6–11 y)	BOT-2 (subtest strength)		Children with DCD had significantly poorer performance compared with TD. No differences between countries found.
Farhat et al, 2015[10]	Cross-sectional	19 DCD, 18 TD (7–9 y)	5JT, THD		Significantly poorer performance in children with DCD
Farhat et al, 2016[31]	RCT	27 DCD, 14 TD (6–10 y)	5JT, THD	3×/wk for 1 h for training group (14 DCD). Non-training group (13 DCD) and TD (14) only get regular classroom activities, physical education classes as scheduled	Children with DCD had lower scores on different outcome measures. Intervention group improved significantly on explosive power, strength, and agility. Control groups (DCD and TD) showed no significant improvement on different outcome measures.
Ferguson et al, 2014[11]	Case control	70 DCD, 70 TD (6–10 y)	FSM, MPST, HHD (elbow flexors, elbow extensors, and knee extensors)		Children with DCD had significantly lower performance on FSM (except overarm throwing and chest pass). No significant difference found on HHD and MPST.
Ferguson et al, 2015[32]	Pre-posttest	22 DCD, 19 TD (6–10 y)	FSM, MPST	Health promotion program within school environment (9 wk)	Children with DCD had significantly lower performance compared with TD. Children with DCD and TD both improved significantly on outcome measures. There was a significant group × time effect on MPST (mean power).

(*Continued*)

Downloaded from https://academic.oup.com/ptj/article/100/10/1825/5868019 by University of Groningen user on 16 October 2020

Table 2.
Continued

Study	Design	Population No. (Age)	Outcome Measures Regarding Strength and Anaerobic Capacity	Intervention	Results
Fong et al, 2005[37]	Cross-sectional exploratory	130 DCD, 117 TD (6–10 y)	Dynamometer Lafayette Manual Muscle Test System		Children with DCD had significantly lower rates of isometric strength (hamstrings and gastrocnemius)
Fransen et al, 2014[38]	Longitudinal study (2 years)	34 relatively low motor competence, 42 relatively average motor competence, and 32 relatively high motor competence (6–10 y) measured with KTK	BOT-2 (sit-up and knee push-up), Eurofit (handgrip strength, standing broad jump, 10 × 5 m sprint)		Children with high motor competence scored better than children with low motor competence. In baseline group 6–8 y, time effect on all measures except handgrip and group effect on all measures. In baseline group 8–10 y, time effect for handgrip and 10 × 5 m sprint, and group effect for all measures except handgrip. Group × time effect for 10 × 5 m sprint.
Haga et al, 2009[12]	Longitudinal study	12 DCD,12 TD (T1 9 y, T2 12 y)	TPF (standing broad jump, jumping 7 m on 1 and 2 feet, throwing a tennis ball, chest pass with medicine ball, climbing wall bars, 10 × 5 m sprint)		Significant differences between groups at T1 and T2. DCD group showed significant improvement over time on 3 of 9 test items: chest pass with medicine ball, climbing wall bars, and reduced Cooper test. TD group showed significant improvement over time on 7 of 9 test items; no significant differences observed in jumping on 2 feet or throwing tennis ball.
Hands et al, 2008[39]	Longitudinal study	18 low motor performance (LMC), 18 TD (5–7 y measured 1×/y for 5 y)	50-m sprint, overhand throw (tennis ball), standing broad jump		Slower speed and lower performance on different fitness outcomes for LMC group. Differences remained similar over time for overhand throw and standing broad jump. Significant group × time effect for sprint indicated decrease of differences between groups.
Hoek van der et al, 2012[14]	Multi-center case control	38 DCD, 38 TD (7–12 y)	HHD, handgrip (Jamar)		Children with DCD had less muscle strength in elbow extension and flexion and knee flexion. Knee extension and handgrip did not differ between groups.
Kanioglou, 2006[15]	Cross-sectional	10 with severe motor coordination problems, 16 with moderate coordination problems, and 125 TD (mean age 10.9 y, SD 0.68)	AAHPERD YFT (pull-up [boys] or flexed arm hang [girls, sit up, long jump, 50-yard dash)		Children with moderate and severe motor problems showed significantly lower scores

(*Continued*)

Downloaded from https://academic.oup.com/ptj/article/100/10/1825/5868019 by University of Groningen user on 16 October 2020

Table 2.
Continued

Study	Design	Population No. (Age)	Outcome Measures Regarding Strength and Anaerobic Capacity	Intervention	Results
Li et al, 2011[40]	Perspective study of 3 years	25 DCD, 25 TD (9–11 y)	Taiwan physical fitness (sit-up, long jump)		Children with DCD performed worse and differences increased over time
Lifshitz et al, 2014[41]	Cross-sectional comparative	22 DCD, 47 TD (6–11 y)	BOT-2 (subtest strength)		Children with DCD had significantly lower performance
Nascimento et al, 2013[42]	Cross-sectional	21 severe DCD, 21 moderate DCD, 21 TD (6–10 y)	Fitnessgram (curl-up, trunk lift, sit-up, push-up)		No significant differences between groups on muscle strength
Raynor et al, 2001[43]	Cross-sectional	20 DCD, 20 TD (6–10 y)	50-m sprint		Children with DCD performed significantly worse
Ruas et al, 2014[44]	Cross-sectional	5 low motor performance, 19 TD (mean age 10.8 y, SD 1.7)	Vertical jump, SLJ		Children with low motor performance had significantly lower scores compared with TD children
Scott et al, 2007[17]	Cross-sectional	155 DCD, 106 TD (4–12 y)	20-m sprint, jump-and-reach test, 1-k medicine ball throw, curl-ups		20-m sprint, jump-and-reach test, and medicine ball throw significantly worse in children with DCD
Smits-Engelsman et al, 2017[33]	Pre-post experimental design	18 DCD, 18 TD (6–10 y)	FSM (lower extremities), 10 × 5 m sprint and slalom, BOT-2 (subtest running speed and agility)	20 min active gaming on balance board, 2×/wk for 5 wk	Children with DCD had lower rates of functional strength, agility, and a trend on sprint tests. After intervention, functional strength and anaerobic fitness improved in TD and DCD children.
Tsiotra et al, 2009[45]	Cross-sectional	12 DCD, 165 TD (12 y)	Vertical jump, HHD,[b] 40-m speed test		Vertical jump, hand strength, and 40-m sprint were significantly poorer in DCD compared with TD children

[a] AAHPERD YFT = American Alliance testing for Health, Physical Education, Recreation and Dance for Youth; BOT-2 = Bruininks-Oseretsky Test of Motor Proficiency-Second Edition; DCD = developmental coordination disorder; FSM = functional strength measurement; HHD = hand-held dynamometer; 5JT = five-jump test; KTK = Körper Coordination Test for Children; LMC = low motor performance; MPST = muscle power sprint test; SLJ = standing long jump; TD = typically developing; THD = triple-hop distance; TPF = test of physical fitness.

[b] T.K.K.5101, Takei Scientific Instruments, Tokyo, Japan.

Downloaded from https://academic.oup.com/ptj/article/100/10/1825/5868019 by University of Groningen user on 16 October 2020

Table 3.
Methodological Qualities of Different Study Designs to Examine Different Psychometric Properties According to Risk of Bias Checklist of COSMIN[a]

	Population No., Y, Country	Measure	Description of Tests for Strength and Anaerobic Capacity	Psychometric Properties	COSMIN Score
Aertssen et al, 2016b[46]	474 TD, 4–10 y, the Netherlands	FSM	Overarm and underarm throwing: throwing distance of heavy sandbag as far as possible. SLJ: jump as far as possible. Lateral step-up: on first step of stairs with 2 fingers against wall (no. of repetitions in 30 sec). Chest pass: sit with back against wall and push heavy bag from chest. Sit-to-stand: stand up and sit down from chair (no. repetitions in 30 sec). Lifting a box: lift box filled with weights and put it on another box (no. of repetitions in 30 sec). Stair climbing: run up and down stairs (no. of steps in 30 sec).	Reliability Measurement error Content validity Structural validity Internal consistency Construct validity	Adequate Adequate Doubtful Very good Very good Very good
Ayán Pérez et al, 2014[47]	120 TD, mean age, 48.60 (SD 9.94) mo, Spain	Bent knee push-up and HHD (Takei)	Bent knee push-up: in a straight line from head to knee pushing up from 90° flexion to extension (no. of correct push-ups). HHD: hand grip.	Reliability Construct validity	Adequate Very good
Beld et al, 2006[48]	64 TD, 4–11 y, the Netherlands	HHD (MicroFet2)	A "make" test was used where HHD was held stationary and child pushed with an isometric contraction in different directions: shoulder extensors and abductors; elbow extensors and flexors; wrist extensors; hip flexors, extensors, and abductors; knee flexors and extensors; ankle dorsiflexors.	Reliability Measurement error Construct validity	Very good Very good Very good
Bongers et al, 2015[49]	65 TD, 6–18 y, the Netherlands	Pediatric RAST	6 × 15-m sprint with 10-sec rest between	Criterion validity	Very good
Bonney et al, 2018b[50]	388 children, 6–16 y; varying groups of 59–86–100 TD and 60–110 DCD per test, South Africa	MPST, 10 × 5 m sprint, 10 × 5 m slalom sprint, shuttle run item of BOT-2	MPST: 6 × 15-m sprint with 10-sec rest between Shuttle run item BOT-2: 15.34-m run picking a block and run back.	Reliability Measurement error Content validity Construct validity Responsiveness	Very good Very good Adequate Very good Doubtful
Brown et al, 2019[51]	123 TD, 8–12 y, Australia	BOT-2 (short form)	Strength subtest: sit-up and knee push-up	Structural validity	Adequate
Cotten et al, 1990[52]	363 TD, 5–11 y, USA	Modified pull-up test	Child placed with shoulders beneath bar 1–2 in beyond reach. Elastic band positioned beneath bar. Child instructed to keep body straight and pull up until chin above elastic band.	Reliability	Doubtful
Davis et al, 2008[53]	105 TD, 5–6 y, USA	Medicine ball throw test (power med-ball) with a weight of 2 lb.; modified pull-up	Medicine ball throw: sitting with back against wall, lift medicine ball to chest and throw as far as possible. Modified pull-up: elastic band positioned beneath bar and child instructed to keep body straight and pull up until chin above elastic band.	Reliability Content validity Criterion validity	Adequate Inadequate Very good

(*Continued*)

Downloaded from https://academic.oup.com/ptj/article/100/10/1825/5868019 by University of Groningen user on 16 October 2020

Table 3.
Continued

	Population No., Y, Country	Measure	Description of Tests for Strength and Anaerobic Capacity	Psychometric Properties	COSMIN Score
Douma-van Riet et al, 2012[54]	379 TD, 6–12 y, the Netherlands	MPST	6 × 15-m sprints with 10-sec rest between	Reliability Measurement error Construct validity	Very good Very good Very good
Duncan et al, 2017[55]	27 TD, 7–10 y, UK	Resistance Training Skills Battery	RTSBc consists of 6 RT skills (bodyweight squat, push-up, step-up, suspended row, standing overhead press, and front support with chest touches).	Construct validity	Very good
Engelman et al, 1991[56]	470 TD, 8–17 y, Greece	Traditional pull-up and modified pull-up	Traditional pull-up: palms facing away from body, pull up. Modified pull-up: elastic band positioned beneath bar and child instructed keep body straight and pull up until chin is above elastic band.	Reliability Construct validity	Doubtful Doubtful
Erbauch et al, 1990[57]	26 TD, mean age 8.3 (SD 1 y), USA	Sit-up and modified pull-up	Minimal description of tests available. Sit-up: no. in 1 min. Modified pull-up: as many as possible.	Reliability Measurement error	Doubtful Doubtful
España-Romero et al, 2010a[58]	66 TD, 12–16 y, Spain	Jamar Hydraulic dynamometer, DynEx electronic hand dynamometer, TKK digital hand dynamometer	Two performances in standing position with each dynamometer, 1 with elbow extended and 1 with elbow in 90° flexion, squeeze for at least 2 sec as hard as possible.	Reliability Measurement error Content validity	Very good Inadequate Inadequate
España-Romero et al, 2010b[59]	138 TD, 6–18 y, Spain	ALPHA health related fitness test (Hand Dynamometry with adjustable grip, TKK 5101 Grip D), standing long jump	Hand Dynamometry: elbow fully extended, squeeze gradually and continuously for at least 2 sec. SLJ: jump as far as possible.	Reliability Measurement error Content validity	Adequate Inadequate Adequate
Fernandez-Santos et al, 2015[60]	368 TD, 6–12 y, Spain	SLJ, squat jump, countermovement jump, and Abalakov jump	SLJ: jump as far as possible. Squat jump: knees bent to 90° and then jump vertically as high as possible. Countermovement jump: stand with extended knees, bend knees to 90° and jump vertically as high as possible with hands on hips. Abalakov jump: squat down until knees at a 90° angle, swing arms back behind body, swing arms forward and jump as high as possible.	Reliability Measurement error Criterion validity	Adequate Very good Very good
Fernandez-Santos et al, 2016[61]	180 TD, 6–12 y, Spain	Hand Dynamometry (TKK 5001 Grip-A; Takey, Tokyo, Japan), basketball throw, push-ups, 1RM bench press test	Hand Dynamometry: elbow in full extension, squeeze as hard as possible for at least 2 sec. Basketball throw: throw ball from behind head as far as possible. Push-up: push up off floor until elbows straight while keeping legs and back straight, as many times as possible.	Reliability Measurement error Content validity Criterion validity	Adequate Very good Very good Very good

(*Continued*)

Downloaded from https://academic.oup.com/ptj/article/100/10/1825/5868019 by University of Groningen user on 16 October 2020

Table 3.
Continued

	Population No., Y, Country	Measure	Description of Tests for Strength and Anaerobic Capacity	Psychometric Properties	COSMIN Score
Fjørtoft et al, 2011[62]	105 TD, 5–12 y, Norway	Test battery	Standing broad jump: jump as far as possible Jumping 7 m on 2 feet and 1 foot: no. of jumps/hops needed to cover 7 m. Throwing a tennis ball with 1 hand: stands with contralateral foot in front of ipsilateral foot. Pushing a medicine ball: stand with 2 feet next to each other and push ball. Climbing wall bars: climbing up wall bars, crossing over 2 columns and climbing down as fast as possible 10 x 5 m shuttle run: run 10 times 5 m without a rest. 20-m sprint.	Reliability Measurement error Internal consistency Construct validity	Adequate Adequate Very good Doubtful
Furzer et al, 2018[63]	21 TD, 19 low motor performance, 6–12 y, Australia	Resistance Training Skills Battery	RTSBc consists of 6 RT skills (bodyweight squat, push-up, step-up, suspended row, standing overhead press, and front support with chest touches).	Reliability Measurement error Structural validity Internal consistency Construct validity	Very good Inadequate Doubtful Very good Adequate
Gerodimos et al, 2013[64]	54 TD, 9–15 y, Greece	Jamar hydraulic dynamometer	Sitting position, feet supported, shoulders adducted and neutrally rotated, elbow flexed at 90° with forearm in neutral and wrist between 0 and 30° of extension with tested arm positioned on a table to support weight of dynamometer.	Reliability Measurement error	Very good Very good
Gajdosik et al, 2005[65]	15 TD, 4 y, USA	HHD (Lafayette)	A "make" test was used where HHD held stationary and child pushed with an isometric contraction until a constant force was recorded for 3 sec for elbow flexion and extension, shoulder flexion, knee extension and flexion.	Reliability	Very good
Hébert et al, 2011[66]	74 TD, 4–7.5 y, Canada	HHD (Chatillon push-pull)	A "make" test was used with stationary HHD held and child pushed with maximal force for: shoulder lateral rotators and abductors; elbow extensors and flexors; hip flexors, extensors, and abductors; knee flexors and knee extensors; ankle plantar flexors and dorsiflexors.	Reliability Measurement error Content validity Criterion validity	Very good Very good Adequate Very good
King Dowling et al, 2017[67]	393 TD, 3–5 y, Canada	SLJ and BOT-2 item shuttle run	SLJ: jump as far as possible. BOT-2 shuttle run item: sprint 15.34 m, pick up a block and run back.	Criterion validity	Very good
Latorre Román et al, 2015[68]	553 TD, 3–6 y, Spain	Two items of fitness test battery	Standing broad jump: jump as far as possible 20-m sprint.	Reliability Measurement error Content validity Construct validity	Adequate Adequate Doubtful Very good
Latorre-Román et al, 2017[69]	3555 TD, 3–6 y, Spain	SLJ	SLJ: jump as far as possible.	Reliability	Adequate
Lucas et al, 2013[70]	30 TD, 7–9 y, Australia	BOT-2 (short form)	Strength subtest: sit-up and knee push-up.	Reliability	Inadequate

(*Continued*)

Downloaded from https://academic.oup.com/ptj/article/100/10/1825/5868019 by University of Groningen user on 16 October 2020

Table 3.
Continued

	Population No., Y, Country	Measure	Description of Tests for Strength and Anaerobic Capacity	Psychometric Properties	COSMIN Score
Macfarlane et al, 2008[71]	17 TD, 6–8 y, USA	HHD (Microfet II)	"Make-test" for knee extension and flexion; hip abduction, adduction, extension, and flexion.	Reliability Content validity	Very good Adequate
Marmis et al, 2013[72]	2060 TD, 9–18 y, USA	AAPHER Youth Fitness Test	Standing broad jump: simultaneously extend knees and swing arms forward to jump as far as possible. Softball throw: throw as far as possible. 50-y run.	Reliability	Doubtful
Ties Molenaar et al, 2008[73]	104 TD, 4–12 y, the Netherlands	Lode dynamometer and Martin vigorimeter	Handgrip strength: measured in sitting position, elbow in 90° flexion and wrist in neutral position	Reliability Measurement error	Inadequate Inadequate
Morrow et al, 2010[74]	1010 TD, 8–17 y, Texas	Fitnessgram	Curl-up: as many as possible up to max of 75. Flexed arm hang: as long as possible. Push-up: as many as possible. Trunk lift: lift upper body off floor and hold position as long as possible.	Reliability Content validity Criterion validity	Very good Adequate Inadequate
O'Connell et al, 2004[75]	69 TD, 4–12 y, Texas	Back extension endurance test	J-Tech Onsite Isometric dynamometer (isometric). Back extensor endurance test: in flexed position with legs and hips lying in 45°, extend trunk until trunk is parallel to floor.	Reliability Construct validity	Doubtful Inadequate
Ramírez-Vélez et al, 2015[76]	229 TD, 9–18 y, Colombia	ALPHA health-related fitness test battery	Standing broad jump: jump as far as possible landing with feet together. Vertical jump: jump as high as possible. Handgrip dynamometer (T-18 TKKSMEDLY II): handgrip, squeeze for 3–5 sec, measured standing with elbows extended. 4 × 10 m sprint: 4×10-m sprint without rest.	Reliability Measurement error Content validity	Inadequate Very good Doubtful
Steenman et al, 2016[77]	683 TD, 6–18 y, the Netherlands	MPST	6 sprints of 15 m with 10-sec rest between sprints	Reliability Measurement error Content validity	Very good Very good Adequate
Vanhelst et al, 2014[78]	174 TD, 8.2–16.2 y, France	BOUGE health-related physical fitness battery	20/30/50-m sprint tests. Basketball throw: throw ball with 2 hands from behind head as far as possible. SLJ: jump as far as possible with feet together.	Reliability Measurement error Content validity	Very good Very good Adequate
Yin L, et al, 2018[79]	240 TD, 10–12 y, China	SMST	Hand-grip with dynamometer (Hui Hai Electronics) in standing position with extended elbow. Knee bent push-up: in straight line from head to knee pushing up from 90° flexion of elbows to extension for 1 min. Sit-up: from lying position to sitting touching outer sides of 2 bended knees for 1 min. SLJ: jump as far as possible with feet naturally apart.	Construct validity Criterion validity	Very good Very good

[a] The qualitative COSMIN scores of the study designs range from very good, adequate, and doubtful to inadequate. 1RM = 1-repetition maximum; BOT-2 = Bruininks-Oseretsky Test of Motor Proficiency-Second Edition; COSMIN = Consensus-Based Standards for the Selection of Health Measurement Instruments; DCD = developmental coordination disorder; FSM = functional strength measurement; HHD = hand-held dynamometer; MPST = muscle power sprint test; RAST = repeated anaerobic sprint test; RTSBc = Resistance Training Skills Battery for Children; SLJ = standing long jump; SMST = simple muscle strength test; TD = typically developing.

Downloaded from https://academic.oup.com/ptj/article/100/10/1825/5868019 by University of Groningen user on 16 October 2020

Table 4.
Reliability and ME and Scores on CPP of Different Measures[a]

Measure	Study	Methodological Quality of Study Design	Reliability	CPP	Measurement Error	CPP
DCD						
Anaerobic sprint tests						
MPST/Pediatric RAST	Bonney et al, 2018b[50]	Very good (R + ME)	Test retest: time ICC .76, mean power ICC .50, peak power ICC .23	–	Time: SEM 0.3 sec, SDD 1 sec mean power: SEM 92.2 Watt (31% of mean score), SDD 254.8 Watt peak power: SEM 346.2 Watt (71% of mean score), SDD 956.9 Watt	–
10 × 5 m sprint	Bonney et al, 2018b[50]	Very good (R + ME)	Test retest: ICC .92	+	SEM 1.2 sec (5.3% of mean score), SDD 3.4 sec	?
10 × 5 m slalom	Bonney et al, 2018b[50]	Very good (R + ME)	Test retest: ICC .92	+	SEM 1.4 sec (6.5% of mean score) SDD 3.9 sec	?
BOT2-SR item	Bonney et al, 2018b[50]	Very good (R + ME)	Test retest: ICC .89	+	SEM 0.7 sec (7% of mean score) SDD 1.9 sec	?
TD						
Isometric strength						
HHD	Hebert et al, 2011[66]	Very good (R + ME)	Intrarater: ICC .79–.98 Interrater ICC .67–.96	+	SEM: 0.5–4.9 Newton	?
	Macfarlane et al, 2008[71]	Very good (R + ME)	Test retest: .82 to .91	+		
	Beld et al, 2006[48]	Very good (R + ME)	Test retest: ICC .83–.95	+	SEM ranged from 3.3 to 12.2 Newton	?
Hand dynamometer	Fernandez-Santos et al, 2016[61]	Adequate (R) very good (ME)	Inter-trial: ICC .98	+	6% error SEE 2.7 kg systematic error near 0	?
	España-Romero et al, 2010[58]	Adequate (R) Inadequate (ME)	Bland-Altman made, reliability tested by inter-trial differences calculated through observed systematic bias	?	Systematic bias 0.02 kg LoA 1.57 kg	?
	Ayán Pérez et al, 2014[47]	Adequate (R)	Test retest: ICC .84–.92	+		
Jamar hydraulic dynamometer	Gerodimos et al, 2013[64]	Very good (R + ME)	Inter-session and intra-session: ICC .87–.99	+	SEM 0.88–1.54 kg	?
	España-Romero et al, 2010[58]	Adequate (R) Inadequate (ME)	Bland-Altman made, reliability tested by inter-trial differences calculated through observed systematic bias	?	Observed systematic bias 0.23 kg LoA 1.20 kg	?
DynEx electronic hand dynamometer	España-Romero et al, 2010[58]	Adequate (R) Inadequate (ME)	Bland-Altman made, reliability tested by inter-trial differences calculated through observed systematic bias	?	Observed systematic bias 0.26 kg LoA 1.42 kg	?
Lafayette Manual Muscle Testing	Gajdosik et al, 2005[65]	Very good (R)	Test retest: ICC .54–.94	+		

(*Continued*)

Downloaded from https://academic.oup.com/ptj/article/100/10/1825/5868019 by University of Groningen user on 16 October 2020

Table 4.
Continued

Measure	Study	Methodological Quality of Study Design	Reliability	CPP	Measurement Error	CPP
Muscle endurance						
Back extensor	O'Connell et al, 2004[75]	Doubtful (R)	Test retest: r .55	?		
Pull-up	Engelman et al, 1991[56]	Doubtful (R)	Test retest: ICC > .91	+		
Modified pull-up	Erbaugh et al, 1990[57]	Doubtful (R + ME)	Test retest: ICC r .52	?	41% error variance (participants × trials)	?
	Cotten et al, 1990[52]	Doubtful (R)	Test retest: ICC r .72–.95	?		
Push-up test	Fernandez-Santos et al, 2016[61]	Adequate (R) Very good (ME)	Inter-trial: ICC .91	+	5% error SEE 2 repetitions systematic error near 0	?
Bent knee push-up	Ayán Pérez et al, 2014[47]	Adequate (R)	Test retest: ICC .70–.85	+		
Sit-up	Erbaugh et al, 1990[57]	Doubtful (R + ME)	Test retest: r .83	?	28% error variance (participants × trials)	?
Muscle power						
SLJ/standing broad jump	Fernandez-Santos et al, 2015[60]	Adequate(R) Very good (ME)	Test retest: ICC .95	+	Systematic error near 0	?
	Latorre-Román et al, 2017[69]	Adequate (R)	Test retest: ICC .91	+		
Medicine ball pushing	Davis et al, 2008[53]	Adequate (R)	Test retest: ICC .88	+		
Basketball throw	Fernandez-Santos et al, 2016[61]	Adequate(R) Very good (ME)	Test retest: ICC .98	+	7% error, SEE 0.62 systematic error near 0	?
Anaerobic sprint tests						
MPST/Pediatric RAST	Douma-van Riet et al, 2012[54]	Very good (R + ME)	Test retest: peak power ICC .98, mean power ICC .98	+	LoA −25% to 22% for mean power	?
	Steenman et al, 2016[77]	Very good (R + ME)	Test retest: ICC .90 Inter rater: ICC .97	+	LoA −16.6 to +16.8 Watt	?
	Bonney et al, 2018b[50]	Very good (R + ME)	Test retest: ICC .70–.91	+	Mean power: systematic error − 15.43 Watt, LoA 134.51/−165.38 Watt, SEM 16.8% of the mean score peak power: systematic error −32.01 Watt, LoA 383.64/−447.67, SEM 35.6% of mean score	?
Test battery						
FSM	Aertssen et al, 2016b[46]	Adequate (R + ME)	Test retest: ICC .77–.94	+	SEM clusters 4–6 y 05–0.69 SS and 0.83 SS for total score. SEM clusters 7–10 0.59–0.97 SS and for total score 1.01 SS. SDC 4–6 y 1.39–1.92 SS and for total score 2.33. SDC 7–10 y 1.63–2.69 SS and for total score 2.80 SS	?

(*Continued*)

Downloaded from https://academic.oup.com/ptj/article/100/10/1825/5868019 by University of Groningen user on 16 October 2020

Table 4.
Continued

Measure	Study	Methodological Quality of Study Design	Reliability	CPP	Measurement Error	CPP
Test battery	Fjørtoft et al. 2011[62]	Adequate (R + ME)	Test retest: ICC .54–.92	+	Measurement error 0.26 for total test (z-score)	?
ALPHA health-related fitness test battery	Ramírez-Vélezet al, 2015[76]	Inadequate (R) Very good (ME)	Test retest: no significant differences between T1 and T2	?	Bland Altman plots showed small ME	?
	España-Romero et al, 2010b[59]		Test retest: Long jump sign difference in children, but not in adolescents, Handgrip no sign difference between T1 and T2	?	Long jump: 6.3% error SEE 13.32 cm Handgrip: 2.28% error SEE 1.99 kg	?
BOUGE health-related fitness battery	Vanhelst et al, 2016[78]	Very good (R + ME)	Test retest: Sprint test ICC .97, Basketball throw ICC .93, Standing broad jump ICC .93	+	Sprint test: mean difference very near 0 (−0.16 ± 0.35 sec), LoA –0.88 to 0.54 Basketball throw: mean difference very near 0 (−18.3 ± 158.6) m and limits of agreement were −335.6 to 298.9 Standing broad jump: mean difference very near 0 (1.8 ± 31.5 cm), LoA −61.3 to 64.9	?
Fitnessgram	Morrow et al, 2010[74]	Very good (R)	Inter- and intra-rater: teacher-teacher: curl-up Kappa .74, push-up Kappa .48, trunk lift Kappa .72 expert-expert curl-up Kappa .56, push-up Kappa .54, trunk lift Kappa .54	−		
Fitness test battery	Latorre Román et al, 2015[68]	Adequate (R + ME)	Test retest: standing broad jump ICC .91, 20-m sprint test ICC .94 Inter-rater: 20-m sprint test ICC .99	+	Standing broad jump: LoA 25.4/−21.4 cm, mean differences 1.96 ± 11.72 cm 20-m sprint test: LoA 1.06/−1.09 sec, mean differences −0.01 ± 0.54 sec	?
AAPHER	Marmis et al, 2013[72]	Doubtful (R)	Correlations between trial in multi-trial: Long jump r > .73, 50-yard run r > .66, Softball throw r > .86	?		
BOT-2 SF (strength items)	Lucas et al, 2013[70]	Inadequate (R)	Test retest: ICC .26–.31 Inter-rater: ICC .86–.87	+/−	SEM and MDC described for whole BOT-2 SF and not specific for strength items	?
Resistance Training Skills Battery for Children	Furzer et al, 2018[63]	Very good (R) Inadequate (ME)	Test-retest: ICC .95–.99	+	Typical error for RTSQc was small 0.35 (95%CI −0.92 to 0.50), SEM, SDC, MIC not described	?

[a] + = positive rating; ? = indeterminate rating; − = negative rating; AAPHER = American Alliance Testing for Health, Physical Education, Recreation and Dance; BOT-2 = Bruininks-Oseretsky Test of Motor Proficiency-Second Edition; CPP = criteria psychometric properties; DCD = developmental coordination disorder; FSM = functional strength measurement; HHD = hand-held dynamometry; ICC = intraclass correlation coefficient; LoA = limits of agreement; MDC = minimal detectable change; ME = measurement error; MPST = muscle power sprint test; R = reliability; RAST = running-based anaerobic sprint test; RTSQ = resistance training skills quotient for children; SEE = standard error of estimate; SEM = standard error of measurement; SDD = smallest detectable difference; SF = short form; SLJ = standing long jump; SS = standard score; TD = typically developing.

Downloaded from https://academic.oup.com/ptj/article/100/10/1825/5868019 by University of Groningen user on 16 October 2020

Table 5.
Validity and Scores on CPP of Different Measures[a]

Measure	Study	Methodological Quality of Study Design	CNT	CPP	CON	CPP	CRI	CPP
DCD								
Anaerobic sprint tests								
MPST/Pediatric RAST	Bonney et al, 2018b[50]	Adequate (CNT) Very good (CON)	Face validity described of different sprint test used	+	Convergent validity: MPST and 10 × 5 m/10 × 5 m slalom/BOT-2 SR: NS Divergent validity: MPST and 20-m SR: NS	–		
10 × 5 m sprint	Bonney et al, 2018b[50]	Adequate (CNT) Very good (CON)	Face validity described of different sprint test used	+	Convergent validity: 10 × 5 m and 10 × 5 m slalom: r .58 10 × 5 m and BOT-2 item SR: r .37	+		
10 × 5 m slalom	Bonney et al, 2018b[50]	Adequate (CNT) Very good (CON)	Face validity described of different sprint test used	+	Convergent validity: 10 × 5 m slalom and 10 × 5: r .31 10 × 5 m slalom and MPST: NS 10 × 5 m slalom and BOT-2 item SR: NS	+		
BOT-2 item SR	Bonney et al, 2018b[50]	Adequate (CNT) Very good (CON)	Face validity described of different sprint test used	+	Convergent validity: BOT-2 item SR and 10 × 5 m: r .37 BOT-2 item SR and 10 × 5 m slalom: NS BOT-2 item SR and MPST: NS	+		
TD								
Isometric strength								
HHD	Hebert et al, 2011[66]	Adequate (CNT) Very good (CRI)	No adaptions must be made in instructions, positions of HHD, or no. of tests for a particular muscle group Also no reports of pain or discomfort during testing	+			Cybex: ICC .48–.94	+
	Macfarlane et al, 2008[71]	Adequate (CNT)	Collecting reference measures, with cut-off values for normal and below-normal strength	+				
	Beld et al, 2006[48]	Very good (CON)			Known group: AUCs ranged from .66 to .88 (muscle biopsy, difference myopathy, and TD)	+		

(*Continued*)

Downloaded from https://academic.oup.com/ptj/article/100/10/1825/5868019 by University of Groningen user on 16 October 2020

Table 5.
Continued

Measure	Study	Methodological Quality of Study Design	CNT	CPP	CON	CPP	CRI	CPP
Hand dynamometer	Fernandez-Santos et al, 2016[61]	Very good (CNT + CRI)	Participants learned protocol and any questions were answered. For comparison test of 1RM bench press, participants were asked how difficult it was to lift weight; depending on answer, weight increased	?			Hand dynamometer-1RM bench press test: r .79	+
	España-Romero et al, 2010a[58]	Inadequate (CRI)					Known weight: systematic bias 0.49 kg, LoA 1.32 kg	?
Jamar hydraulic dynamometer	España-Romero et al, 2010a[58]	Inadequate (CRI)					Known weight: systematic bias −1.92 kg, LoA 1.92 kg	?
DynEx electronic hand dynamometer	España-Romero et al, 2010a[58]	Inadequate (CRI)					Known weight: systematic bias −1.43 kg, LoA 3.56 kg	?
Muscle endurance								
Back extensor	O'Connell et al, 2004[75]	Inadequate (CON)			A J-Tech Onsite Isometric dynamometer and back extension: NS	–		
Pull-up	Engelman et al, 1991[56]	Doubtful (CON)			Pull-up and modified pull-up: r .49–64	?		
Push-up test	Fernandez-Santos et al, 2016[61]	Very good (CNT + CRI)	Participants learned protocol and any questions were answered. For 1RM bench press, participants were asked how difficult it was to lift weight; depending on answer, weight increased	?			Push-up- 1RM bench press: r .21	–
Bent knee push-up	Ayán Pérez et al, 2014[47]	Very good (CON)			Bent knee push-up and handgrip: r .52–.82.	+		

(*Continued*)

Downloaded from https://academic.oup.com/ptj/article/100/10/1825/5868019 by University of Groningen user on 16 October 2020

Table 5.
Continued

Measure	Study	Methodological Quality of Study Design	CNT	CPP	CON	CPP	CRI	CPP
Muscle power								
SLJ/standing broad jump	Fernandez-Santos et al, 2015[60]	Very good (CRI)					SLJ- 1RM leg extension test: r .40, SLJ-standardized for weight: r .79 SLJ-CMJ: r .74 SLJ-SJ: r .73, SLJ-AJ: r .78	+
	King-Dowling et al, 2017[67]	Very good (CRI)					SLJ-Wingate: r .64	–
Medicine ball pushing	Davis et al, 2008[53]	Inadequate (CNT) Very good (CRI)	Medicine ball throw positively related with height (r .34) and weight (r .34). Significant difference between 5-y-old and 6-y-old groups, which supports correlational and known-difference evidence of validity for Medicine ball throw test	?			Medicine ball pushing-modified pull-up test: NS	–
Basketball throw	Fernandez-Santos et al, 2016[61]	Very good (CNT + CRI)	Participants learned protocol and any questions were answered. Participants were asked how difficult it was to lift weight; depending on answer, weight increased	?			Basketball throw- 1RM bench press test: r .69	–
Anaerobic sprint tests								
MPST/Pediatric RAST	Douma-van riet et al, 2012[54]	Very good (CON)			Known group: high sport participation had significant higher mean power compared with moderate or low sport participation. Boys significantly higher mean and peak power than girls ($P < .01$)	+		

(*Continued*)

Downloaded from https://academic.oup.com/ptj/article/100/10/1825/5868019 by University of Groningen user on 16 October 2020

Table 5.
Continued

Measure	Study	Methodological Quality of Study Design	CNT	CPP	CON	CPP	CRI	CPP
	Steenman et al, 2016[77]	Adequate (CNT)	Collect more reference values in older age group, after which groups of former study included and transformed into height-related normative reference curves	+				
	Bonney et al, 2018b[50]	Adequate (CNT) Very good (CON)	Face validity described of different sprint test used	+	Convergent validity: MPST and 10 × 5 /10 × 5 m slalom/BOT-2 item SR: NS Divergent validity: MPST and 20-m SR: r .38–.48	+		
	Bongers et al, 2015[49]	Very good (CRI)					RAST-Wingate: r .86–.91. Age children-RAST/Wingate: r .85–.90	+
10 × 5 m sprint	Bonney et al, 2018b[50]	Adequate (CNT) Very good (CON)	Face validity described of different sprint test used	+	10 × 5 m and 10 × 5 m slalom: r .31 10 × 5 m and MPST: NS 10 × 5 m and BOT-2 item SR: r .52	+		
10 × 5 m slalom	Bonney et al, 2018b[50]	Adequate (CNT) Very good (CON)	Face validity described of different sprint test used	+	10 × 5 slalom and 10 × 5 m: r .31 10x5m slalom and MPST: NS 10 × 5 m slalom and BOT-2 item SR: NS	+		
BOT-2 item SR	Bonney et al, 2018b[50]	Adequate (CNT) Very good (CON)	Face validity described of different sprint test used	+	BOT-2 item SR and 10 × 5: r .52 BOT-2 item SR and 10 × 5 m slalom: NS BOT-2 item SR and MPST: NS	+		
	King Dowling et al, 2017[67]	Very good (CRI)					BOT-2 item SR-Wingate: r .68	–
ICF level activity test battery								
FSM	Aertssen et al, 2016b[46]	Doubtful (CNT) Very good (CON)	Items and construct behind items described; expert panel was used	+	Convergent validity: FSM and HHD: r .42–.74 Discriminant validity: FSM and MABC-2: r .23–.39	+		
Test battery	Fjørtoft et al, 2011[62]	Doubtful (CON)			Convergent validity: test battery and evaluation physical fitness by PE teacher: r .90–.93	+		

(*Continued*)

Downloaded from https://academic.oup.com/ptj/article/100/10/1825/5868019 by University of Groningen user on 16 October 2020

Table 5.
Continued

Measure	Study	Methodological Quality of Study Design	CNT	CPP	CON	CPP	CRI	CPP
ALPHA health-related fitness test battery	Ramírez-Vélezet al, 2015[76]	Doubtful (CNT)	Items and construct behind items described	+				
	España-Romero et al, 2010b[59]	Adequate (CNT)	Feasibility investigated on clothes, understood instructions, rejection, motivation; for PE teachers on facility: easy to administer, previous experience, and time to prepare and administer. Acceptable level of feasibility considered when items were "positively" answered in at least 95% of cases	+				
Bouge health-related fitness battery	Vanhelst et al, 2016[78]	Adequate (CNT)	Feasibility for PE teachers: all reported good feasibility in administering tests, costs, and in/outdoor possible	+				
Fitness test battery	Latorre Román et al, 2015[68]	Doubtful (CNT) Very good (CON)	Tests used were safe, easy to perform, very acceptable, and understandable by children. Test performance increases with age	+	Convergent validity: Standing broad jump and 20-m sprint: r .51	+		
SMST	Yin et al, 2018[79]	Very good (CON+CRI)			Correlations between different items of SMST: r .23–.77	+	SMST- Biodex: r .42–.81	−
Resistance Training Skills Battery for Children	Furzer et al, 2018[63]	Adequate (CON)			RTSQc and 5RM (chestpress, leg press, and pull-down) strength scores r = 0.61, $P < .001$	+		
	Duncan et al, 2017[55]	Very good (CON)			Children who scored higher or lower on RTSBc did not significantly differ on measures of muscular strength	?		

[a] + = positive rating; ? = indeterminate rating; − = negative rating; AJ = Abalakov jump; BOT-2 SR and BOT-2 = Bruininks-Oseretsky Test of Motor Proficiency-Second Edition; CMJ = counter movement jump; CNT = content validity; CON = construct validity; CPP = criteria psychometric properties; CRI = criterion validity; DCD = developmental coordination disorder; FSM = functional strength measurement; HHD = hand-held dynamometer; ICC = intraclass correlation coefficient; ICF = International Classification of Functioning, Disability and Health; LoA = limits of agreement; NS = not significant; MPST = muscle power sprint test; PE = physical education; RAST = running-based anaerobic sprint test; RM = repetition maximum; SLJ = standing long jump; SJ = squat jump; SMST = simple muscle strength test; SR = shuttle run; TD = typically developing.

Downloaded from https://academic.oup.com/ptj/article/100/10/1825/5868019 by University of Groningen user on 16 October 2020

Table 6.
Overview of Psychometric Criteria Found in Studies (With Adequate or Very Good Methodological Quality) for Anaerobic Capacity and Muscle Strength Measures in TD Children and Children With DCD[a]

Measure	CNT	Structural Validity	Internal Consistency	Reliability	ME	CON	CRI
Isometric strength							
HHD	+	NA	NA	+	?	?	+
Hand dynamometer	?	NA	NA	+	?	NA	+
Jamar hydraulic dynamometer	NA	NA	NA	+	NA	NA	?
DynEx electronic hand dynamometer	NA	NA	NA	NA	NA	NA	?
Lafayette Manual Muscle Testing	NA	NA	NA	+	NA	NA	NA
Muscle endurance							
Back extensor	NA	NA	+	?	NA	+/−	NA
Pull-up	NA	NA	NA	+	NA	?	NA
Modified pull-up	NA	NA	NA	?	?	NA	NA
Push-up	?	NA	NA	+	?	NA	−
Bent knee push-up	NA	NA	NA	+	NA	+	NA
Sit-up	NA	NA	NA	+	NA	+	NA
Muscle power							
Standing long jump/standing broad jump	NA	NA	NA	+	?	NA	+/−
Medicine ball pushing	NA	NA	NA	+	NA	NA	−
Basketball throw	?	NA	NA	+	?	NA	−
Anaerobic sprint tests							
MPST/RAST	+	NA	NA	+ (DCD−)	?	+	+
10 × 5 m sprint	+	NA	NA	DCD+	DCD?	+	NA
10 × 5 m slalom	+	NA	NA	DCD+	DCD?	+	NA
BOT-2 item SR	+	NA	NA	DCD+	DCD?	+	NA
Test battery							
FSM	+	+	+	+	?	+	NA
Test battery	NA	NA	+	+	?	+	NA
ALPHA health-related fitness test battery	+	NA	NA	?	?	NA	NA
BOUGE health-related fitness battery	+	NA	NA	+	?	NA	NA
Fitnessgram	NA	NA	NA	−	NA	NA	?
Fitness test battery	+	NA	NA	+	?	+	NA
AAPHER	NA	NA	NA	?	NA	NA	NA
BOT-2 SF (strength items)	NA	−	NA	+/−	NA	NA	NA
SMST	NA	NA	NA	NA	NA	+	−
Resistance Training Skills Battery for Children	NA	+	+	+	?	?	NA

[a] + = positive psychometric criteria, − = negative psychometric criteria; ? = indeterminate psychometric criteria; AAPHER = American Alliance Testing for Health, Physical Education, Recreation and Dance; BOT-2 = Bruininks-Oseretsky Test of Motor Proficiency-Second Edition; CNT = content validity; CON = construct validity; CRI = criterion validity; DCD = developmental coordination disorder; FSM = functional strength measurement; HHD = hand-held dynamometer; ME = measurement error; MPST = muscle power sprint test; NA = not available; RAST = running-based anaerobic sprint test; SMST = simple muscle strength test; SR = shuttle run; TD = typically developing.

Downloaded from https://academic.oup.com/ptj/article/100/10/1825/5868019 by University of Groningen user on 16 October 2020

also not been well investigated in children with TD. Specifically, information regarding validity and responsiveness is lacking. If tests are psychometrically sound for TD, this may or may not be the case for children with DCD. So even if this information were available for children with TD, the results in children with DCD may be different. Therefore, we strongly suggest research is performed with children with DCD as participants on the psychometric properties of test that already can be recommended for children with TD.

Based on the available evidence about measurement properties from well-designed studies, we made the following recommendations regarding field-based tests to assess strength and anaerobic capacity in children with TD and children with DCD.

Isometric Strength

Overall, one can conclude that dynamometry is a reliable and valid way to provide quantifiable measurements of the isometric strength of a muscle (group). The reliability and validity of the different dynamometers proved to be sufficient to measure strength in children with TD[80,81] aged between 4 to 17.5 years, preferably using the "make" test of the hand-held dynamometer (HHD).[65,66] The reviewed evidence supported the recommendation by Castro-Pinero et al (2010) to perform the hand-grip test with the elbow extended for manual isometric strength testing in children.[82]

Although research regarding responsiveness was not found, the HHD was used in several intervention studies in children with DCD.[20,22,24] The study of Bonney et al (2018a), which focused on improving the levels of physical fitness, reported significant changes in HHD scores, while the study of Ferguson et al (2013), which focused on skill learning, did not. It could be that intervention following the American College of Sports Medicine (ACSM) strength training guideline, like the neuromotor task training and Wii training by Bonney et al (2018a), lead to better effects in isometrically measured strength. Since information regarding measurement error and responsiveness is missing, the evidence of significant changes should be taken with precaution until it is known whether the HHD is sensitive enough to measure change in children with DCD.

Based on our evaluation of the evidence, we recommend the use of HHD to measure isometric strength in children with TD. This instrument offers an assessment of strength that does not require coordination or agility, which may be appropriate for the use in children with DCD. On the other hand, testing strength within an activity may be more ecologically valid. No evidence is available to allow specific recommendations for the use of HHD in children with DCD.

Muscle Endurance

There are several single test items for muscle endurance used in studies of children with DCD compared with their peers who are TD; in addition, different test batteries include muscle endurance items. The psychometric properties of some of these test items and test batteries were investigated in children with TD, but not in children with DCD. Combining the methodological quality (adequate, good, very good) with the analyses of the psychometric properties of the single test items suggests that push-up and bent knee push-up are reliable tests to measure muscle endurance in children in the age range of 6 to 17 years[47,61] and the construct validity of the bent knee push-up is valid[61] in children with TD between 4 and 12 years.

Items testing muscle endurance are also present in test batteries such as the FSM (lateral step-up, sit-to-stand test, lifting a box, and stair climb test), the test battery (jumping a distance of 7 m on 2 feet and 1 foot and climbing wall bars), the Fitnessgram (curl-up, flexed arm hang, push-up, and trunk lift), Bruininks-Oseretsky Test of Motor Proficiency-Second Edition (BOT-2; sit-up, push-up, wall sit, v-up), the simple muscle strength test (knee bent push-up, sit-up), and Resistance Training Skills Battery for Children (RTSBc) (push-up, step-up, body-weight squat, standing overhead press, front support with chest touches, suspended row). The summarized evidence showed that the FSM (4–10 years), test battery (5–12 years), and RTSBc (6–12 years) are reliable and valid measures to use in children with TD. Importantly, movements with fast concentric and eccentric contractions like push-ups may be more difficult for children with DCD to perform. The question for many tests remains whether they are measuring muscle endurance or the ability to anticipate fast directional changes, which is known to be compromised in children with DCD.[84,86]

Although information about responsiveness is lacking, some test batteries have been used in different intervention studies in children with DCD (FSM, BOT-2 subtest strength), showing significant improvement.[20,22,83] However, it is still unknown whether these tests are sensitive enough to measure significant change after intervention beyond the minimal important change or measurement errors, which requires clinicians and researchers to be careful in drawing conclusions.

Muscle Power

There are several single test items for muscle power used in comparative studies of children with DCD and children with TD. The vertical jump (6–18 years), standing long jump (3–12 years),[59,66,68] and medicine ball throw (5–6 years)[52] were investigated in TD and shown to be reliable and the different jump tests were also valid.[59,66] The validity of the five-jump test, triple-hop distance, and jump-and-reach test were not investigated in children with

Downloaded from https://academic.oup.com/ptj/article/100/10/1825/5868019 by University of Groningen user on 16 October 2020

DCD. However, these are good examples of movements in which not only muscle power but also technique and coordination are important performance components and may therefore have different validity in children with DCD.

Items measuring muscle power are also present in different test batteries such as the FSM[46] (overarm throwing, standing long jump, underarm throwing, and chest pass), the test battery[62] (standing broad jump and pushing a medicine ball), the ALPHA health-related fitness test battery[59,76] (standing long jump and vertical jump), the BOUGE health-related fitness battery[78] (basketball throw and standing long jump), the Fitness test battery battery[68] (standing broad jump and 20-m sprint), and the American Alliance Testing for Health, Physical Education, Recreation and Dance[72] (standing long jump and softball throw). The summarized available evidence showed that the FSM (4–10 years), BOUGE health-related fitness battery (8–16 years), fitness test battery (3–6 years), and test battery (5–12 years) are reliable measures to use in children with TD. The construct of the FSM, test battery, fitness test battery, and simple muscle strength test[79] were shown to be valid in children with TD. So far, only the upper limb section of the FSM, specifically for measuring muscle power, has been recommended by Bieber et al (2016); however, this recommendation was based on only 2 identified studies of children with DCD.[86]

Although information about responsiveness is lacking for all the tests, children with DCD showed improvement after intervention on HHD, standing long jump, 5-jump test, triple-hop distance, muscle power sprint test (MPST), FSM, and BOT-2 subset running speed and agility have been used in different intervention studies.[20–22,30–33] This indicates that these measures may be sensitive to detect change, but again it is essential to draw conclusions carefully.

Sprint Tests

In children with TD aged 6 to 18 years, the MPST and repeated anaerobic sprint test (RAST) have been found to be reliable[50,54,77] and valid.[49,50] Known group validity of the MPST has been investigated in different ways. In the study by Douma et al (2012), high-sport participation showed significantly higher mean power compared with moderate or low sport participation.[54] Other studies compared children with TD with children with DCD. In contrast, most studies showed no significant differences between children with DCD and their peers with TD[11,21,22] except for 1 study that did find significant group differences.[8] This lack of differences was ascribed by the authors to the fact that the data were collected in lower social economic areas in South Africa. In these areas, participation in sports and outdoor activities is more difficult for all children, which may explain the lower performance on the MPST for children with TD, similar to the performance of the children with DCD. One study was found that included psychometrics properties of field-based tests for strength and anaerobic capacity in children with DCD. In this study, the shuttle run test (sub-item running speed and agility of BOT-2) and the 10 m × 5 m (10 × 5 m) sprint tests (straight and slalom) were found to be reliable and valid.[50] However, in children with DCD, no significant or low intraclass correlation coefficients were found on the MSPT/RAST. This was explained by greater variance in performance especially at the first recording of the test-retest trials.[50] Responsiveness was also investigated. The 10 × 5 m sprint and the 10 × 5 m slalom sprint were found to be responsive to change.[50]

The MPST/RAST was considered to be an appropriate test for children with TD. For children with DCD, we recommend the shuttle run test (sub-item of running speed and agility of BOT-2) and the 10 × 5 m sprint tests (straight and slalom) albeit based on the results of only 1 study.

Recommendations

Generally, we can conclude that information regarding the psychometric properties of field-based tests for strength and anaerobic capacity in children with DCD is lacking. It is important to investigate the psychometric properties in a specific target group because attributes of that specific group, such as a lower level of coordination, may influence the outcomes. For children with TD, more information on the psychometric properties of tests is available but still incomplete, as our results show. Particularly, information concerning the validity and responsiveness is often missing. Measurement error is described for some items and most test batteries, but information about the minimal important change is absent. Furthermore, the most frequent age range studied lies between 6 and 10 years. Hence, the psychometric properties of most measures have not yet been tested in children in the full age range for which they are intended, and this should be taken into account when interpreting the results. In future studies, when evaluating the psychometric properties for strength and anaerobic condition in children with DCD, it might be of interest to determine if the severity of the motor coordination problems or co-morbidity influences the validity or reliability.

Based on our review, we are able to make recommendations for the use of strength and anaerobic capacity measures that have good reliability and validity in children who have TD. The following tools have shown to have the best psychometric qualities in children with TD: HHD for isometric strength; bent knee push-up, FSM, test battery, and RTSBc for muscle endurance; and vertical jump, standing long jump, FSM, and fitness test for muscle power.

For children with DCD, only the shuttle run item of the BOT-2 and the 10 × 5 m sprint (straight and slalom) were investigated and are the best choices at this moment.

Downloaded from https://academic.oup.com/ptj/article/100/10/1825/5868019 by University of Groningen user on 16 October 2020

For all the other strength and anaerobic capacity tests, information is lacking. We highlight that it is important to keep this lack of evidence for the validity and reliability of the outcomes for this target group in mind. As pediatric physical therapists need to measure strength and anaerobic fitness in children with DCD in everyday clinical practice, we advise that in the interim the most valid and reliable tools tested in children with TD should be used until more evidence becomes available. However, clinicians should be aware that results obtained using tests that have shown to be valid and reliable in children with TD may not have the same properties in children with DCD; thus, results should be interpreted carefully and, if possible, verified by different sources. The preferred tests suitable for children with DCD should include different types of strength and power but with low requirements for coordination.

This review showed that there is an urgent need to investigate the psychometric properties of the above-mentioned field-based tests in children with DCD, because outcomes of these tests are used in clinical practice for diagnostic and evaluative purposes.

Author Contributions

Concept/idea/research design: W. Aertssen, D. Jelsma, B.C.M. Smits-Engelsman
Writing: W. Aertssen, D. Jelsma, B.C.M. Smits-Engelsman
Data collection: W. Aertssen, D. Jelsma
Data analysis: W. Aertssen, D. Jelsma, B.C.M. Smits-Engelsman
Project management: B.C.M. Smits-Engelsman
Consultation (including review of manuscript before submitting): D. Jelsma

Funding

There is no funding to report for this study.

Systematic Review Registration

This systematic review was not registered in PROSPERO.

Disclosures

The authors completed the ICMJE Form for Disclosure of Potential Conflicts of Interest. W. Aertssen disclosed that she is a developer of the FSM test battery and that she receives personal fees from joint authorship of the FSM. The other authors declared no conflicts of interest.

DOI: 10.1093/ptj/pzaa118

References

1 Holm I, Fredriksen P, Fosdahl M, Vøllestad N. A normative sample of isotonic and isokinetic muscle strength measurements in children 7 to 12 years of age. *Acta Paediatr*. 2008;97:602–607.

2 Rivilis I, Hay J, Cairney J, Klentrou P, Liu J, Faught BE. Physical activity and fitness in children with developmental coordination disorder: a systematic review. *Res Dev Disabil*. 2011;32:894–911.

3 Pate RR, Wang CY, Dowda M, Farrell SW, O'Neill JR. Cardiorespiratory fitness levels among US youth 12 to 19 years of age: findings from the 1999-2002 National Health and Nutrition Examination Survey. *Arch Pediatr Adolesc Med*. 2006;160:1005–1012.

4 Lee IM, Shiroma EJ, Lobelo F, et al. Effect of physical inactivity on major non-communicable diseases worldwide: an analysis of burden of disease and life expectancy. *Lancet*. 2012;380: 219–229.

5 Rowland TW. *Children's Exercise Physiology*. 2nd ed. Champaign, IL, USA: Human Kinetics; 2005.

6 Katzmarzyk PT, Lear SA. Physical activity for obese individuals: a systematic review of effects on chronic disease risk factors. *Obes Rev*. 2012;13:95–105.

7 Blank R, Barnett AL, Cairney J, et al. International clinical practice recommendations on the definition, diagnosis, assessment, intervention, and psychosocial aspects of developmental coordination disorder. *Dev Med Child Neurol*. 2019;61:242–285.

8 Aertssen WFM, Ferguson GD, Smits-Engelsman BCM. Performance on functional strength measurement and muscle power sprint test confirm poor anaerobic capacity in children with developmental coordination disorder. *Res Dev Disabil*. 2016a;59:115–126.

9 Cantell M, Crawford S, Doyle-Bakker PK. Physical fitness, and health indices in children, adolescents and adults with high or low level motor competence. *Hum Mov Sci*. 2008;27:344–362.

10 Farhat F, Hsairi I, Baiti H, et al. Assessment of physical fitness and exercise tolerance in children with developmental coordination disorder. *Res Dev Disabil*. 2015;45–46:210–219.

11 Ferguson GD, Aertssen WFM, Rameckers EEA, Jelsma J, Smits-Engelsman BCM. Physical fitness in children with developmental coordination disorder: measurement matters. *Res Dev Disabil*. 2014;35:1087–1097.

12 Haga M. The relationship between physical fitness and motor competence in children. *Child Care Health Dev*. 2009;34: 329–334.

13 Hands B, Larkin D. Physical fitness differences in children with and without motor learning difficulties. *Eur J Spec Needs Educ*. 2006;21:447–456.

14 Hoek FD, van der Stuive I. Renders-Messelink HA, et al. Health-related physical fitness in Dutch children with developmental coordination disorder. *J Dev Behav Pediatr*. 2012;33:649–655.

15 Kanioglou A. Estimation of physical abilities of children with developmental coordination disorder. *Studies in Physical Culture and Tourism*. 2006;13:25–32.

16 O'Beirne C, Larkin D, Cable T. Coordination problems and anaerobic performance in children. *Adapt Phys Activ Q*. 1994; 11:141–149.

17 Schott N, Alof V, Hultsch MD. Physical fitness in children with developmental coordination disorder. *Res Q Exerc Sport*. 2007; 78:438–450.

18 Cairney J, Hay JA, Faught BE, Wade TJ, Corna LM, Flouris A. Developmental coordination disorder, generalized self-efficacy toward physical activity, and participation in organized and free play activities. *J Pediatr*. 2006;147:515–520.

19 Magalhaes LC, Cardoso AA, Missiuna C. Activities and participation in children with developmental coordination disorder: a systematic review. *Res Dev Disabil*. 2011;32: 1309–1316.

20 Bonney E, Ferguson G, Smits-Engelsman B. The efficacy of two activity-based interventions in adolescents with developmental coordination disorder. *Res Dev Disabil*. 2017a; 71:223–236.

21 Bonney E, Rameckers E, Ferguson G, et al. "Not just another Wii training": a graded Wii protocol to increase physical

Downloaded from https://academic.oup.com/ptj/article/100/10/1825/5868019 by University of Groningen user on 16 October 2020

fitness in adolescent girls with probable developmental coordination disorder-a pilot study. *BMC Pediatr*. 2018a;18:78.

22 Ferguson GD, Jelsma D, Jelsma J, Smits-Engelsman BCM. The efficacy of two task-orientated interventions for children with developmental coordination disorder: neuromotor task training and Nintendo Wii fit training. *Res Dev Disabil*. 2013;34:2449–2461.

23 Hands B, Chivers P, Grace T, McIntyre F. Time for change: fitness and strength can be improved and sustained in adolescents with low motor competence. *Res Dev Disabil*. 2019;84:131–138.

24 Kordi H, Sohrabi M, Saberi Kakhki A, Hossini SRA. The effect of strength training based on process approach intervention on balance of children with developmental coordination disorder. *Arch Argent Pediatr*. 2016;114:526–533.

25 Wilmore JK, Costill DL. *Physiology of Sport and Exercise*. Champaign, IL, USA: Human Kinetics; 2007.

26 Falgairette G, Bedu M, Fellmann N, et al. Evaluation of physical fitness from field tests at high altitude in circumpubertal boys: comparison with laboratory data. *Eur J Appl Physiol Occup Physiol*. 1994;69:36–43.

27 Consensus-Based Standards for the Selection of Health Measurement Instruments (COSMIN). www.cosmin.nl Accessed May 28, 2020.

28 American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders*. 5th ed. Arlington, VA, USA: American Psychiatric Publishing Incorporated; 2013.

29 Terwee CB, Bot SD, de Boer MR, et al. Quality criteria were proposed for measurement properties of health status questionnaires. *J Clin Epidemiol*. 2007;60:34–42.

30 Bonney E, Jelsma LD, Ferguson GD, Smits-Engelsman BCM. Learning better by repetition or variation? Is transfer at odds with task specific training? *PLoS One*. 2017b;12:e0174214.

31 Farhat F, Hsairi I, Baati H, et al. The effect of a motor skills training program in the improvement of practiced and non-practiced tasks performance in children with developmental coordination disorder (DCD). *Hum Mov Sci*. 2016;46:10–22.

32 Ferguson GD, Naidoo N, Smits-Engelsman BC. Health promotion in a low-income primary school: children with and without DCD benefit, but differently. *Phys Occup Ther Pediatr*. 2015;35:147–162.

33 Smits-Engelsman BCM, Jelsma LD, Ferguson GD. The effect of exergames on functional strength, anaerobic fitness, balance and agility in children with and without motor coordination difficulties living in low-income communities. *Hum Mov Sci*. 2017;55:327–337.

34 Beutum MN, Cordier R, Bundy A. Comparing activity patterns, biological, and family factors in children with and without developmental coordination disorder. *Phys Occup Ther Pediatr*. 2013;33:174–185.

35 Cairney J, Missiuna C, Timmons BW, et al. The coordination and activity tracking in CHildren (CATCH) study: rationale and design. *BMC Public Health*. 2015;15:1266.

36 Cermak SA, Katz N, Weintraub N, et al. Participation in physical activity, fitness, and risk for obesity in children with developmental coordination disorder: a cross-cultural study. *Occup Ther Int*. 2015;22:163–173.

37 Fong SS, Ng SS, Guo X, et al. Deficits in lower limb muscle reflex contraction latency and peak force are associated with impairments in postural control and gross motor skills of children with developmental coordination disorder: a cross-sectional study. *Medicine (Baltimore)*. 2015;94: e1785.

38 Fransen J, Deprez D, Pion J, et al. Changes in physical fitness and sports participation among children with different levels of motor competence: a 2-year longitudinal study. *Pediatr Exerc Sci*. 2014;26:11–21.

39 Hands B. Changes in motor skill and fitness measures among children with high and low motor competence: a five-year longitudinal study. *J Sci Med Sport*. 2008;11:155–162.

40 Li YC, Wu SK, Cairney J, Hsieh C-Y. Motor coordination and health-related physical fitness of children with developmental coordination disorder: a three-year follow-up study. *Res Dev Disabil*. 2011;32:2993–3002.

41 Lifshitz N, Raz-Silbiger S, Weintraub N, Steinhart S, Cermak SA, Katz N. Physical fitness and overweight in Israeli children with and without developmental coordination disorder: gender differences. *Res Dev Disabil*. 2014;35:2773–2780.

42 Nascimento RO, Ferreira LF, Goulardins JB, et al. Health-related physical fitness children with severe and moderate developmental coordination disorder. *Res Dev Disabil*. 2013;34:4222–4231.

43 Raynor AJ. Strength, power, and coactivation in children with developmental coordination disorder. *Dev Med Child Neurol*. 2001;43:676–684.

44 Ruas CV. Strength and power in children with low motor performance scores: a descriptive analysis. *BJMB*. 2014;8:1–9.

45 Tsiotra GD, Nevill AM, Lane AM, Koutedakis Y. Physical fitness and developmental coordination disorder in Greek children. *Pediatr Exerc Sci*. 2009;21:186–195.

46 Aertssen WFM, Ferguson GD, Smits-Engelsman BCM. Reliability, structural and construct validity of the functional strength measurement (FSM) in children aged 4-10 years. *Phys Ther*. 2016b;96:888–897.

47 Ayán Pérez C, Cancela JM, Senra I, Quireza E. Validity and reliability of 2 upper-body strength tests for preschool children. *J Strength Cond Res*. 2014;28:3224–3233.

48 Van den Beld WA, van der Sanden GA, Sengers RC, Verbeek ALM, Gabreëls FJM. Validity and reproducibility of hand-held dynamometry in children aged 4-11 years. *J Rehabil Med*. 2006;38:57–64.

49 Bongers BC, Werkman MS, Blokland D, et al. Validity of the pediatric running-based anaerobic sprint test to determine anaerobic performance in healthy children. *Pediatr Exerc Sci*. 2015;27:268–276.

50 Bonney E, Aertssen W, Smits-Engelsman B. Psychometric properties of field-based anaerobic capacity tests in children with developmental coordination disorder. *Disabil Rehabil*. 2018b;6:1–12.

51 Brown T. Structural validity of the Bruininks-Oseretsky Test of Motor Proficiency—second edition brief form (BOT-2-BF). *Res Dev Disabil*. 2019;85:92–103.

52 Cotton DJ. An analysis of the NCYFS II modified pull-up test. *Res Q Exerc Sport*. 1990;61:272–274.

53 Davis KL, Kang M, Boswell BB, DuBose KD, Altman SR, Binkley HM. Validity and reliability of the medicine ball throw for kindergarten children. *J Strength Cond Res*. 2008;22: 1958–1963.

54 Douma-van Riet D, Verschuren O, Jelsma D, Kruitwagen C, Smits-Engelsman B, Takken T. Reference values for the muscle power sprint test in 6-to 12-year-old children. *Pediatr Phys Ther*. 2012;24:327–332.

55 Duncan MJ, Lawson C, Hurst J, Tallis J, Jones V, Eyre ELJ. Construct validity of the Resistance Training Skills Battery in children aged 7-10 years. *J Sports Sci*. 2018;36:1979–1984.

56 Engelman ME, Morrow JR. Reliability and skinfold correlates for traditional and modified pull-ups in children grades 3-5. *Res Q Exerc Sport*. 1991;62:88–91.

57 Erbaugh SJ. Reliability of physical fitness tests administered to young children. *Percept Mot Skills*. 1990;71:1123–1128.

58 España-Romero V, Ortego FB, Vincent Roderiquez G, Artero EG, Rey JP, Ruiz JR. Elbow position affects handgrip strength in adolescents: validity and reliability of Jamar, DynEx, and TKK dynamometers. *Int J Sports Med*. 2010a;31:490–497.

59 España-Romero V, Artero EG, Jimenez-Pavón D, et al. Assessing health-related fitness tests in the school setting: reliability, feasibility and safety; the ALPHA study. *Int J Sports Med*. 2010b;31:490–497.

60 Fernandez Santos JR, Ruiz JR, Cohen DD, Gonzalez-Montesinos JL, Castro-Piñero J. Reliability and

Downloaded from https://academic.oup.com/ptj/article/100/10/1825/5868019 by University of Groningen user on 16 October 2020

validity of tests to assess lower-body muscular power in children. *J Strength Cond Res*. 2015;29:2277–2285.

61 Fernandez Santos JR, Ruiz JR, Gonzalez-Montesinos JL, Castro-Piñero J. Reliability and validity of field-based tests to assess upper-body muscular strength in children aged 6-12 years. *Pediatr Exerc Sci*. 2016;28:331–340.

62 Fjørtoft I, Pedersen AV, Sigmundsson H, Vereijken B. Measuring physical fitness in children who are 5 to 12 years old with a test battery that is functional and easy to administer. *Phys Ther*. 2011;91:1087–1095.

63 Furzer BJ, Bebich-Philip MD, Wright KE, Reid SL, Thornton AL. Reliability and validity of the adapted Resistance Training Skills Battery for Children. *J Sci Med Sport*. 2018;21:822–827.

64 Gerodimos V, Karatrantou K. Reliability of maximal handgrip strength test in pre-pubertal and pubertal wrestlers. *Pediatr Exerc Sci*. 2013;25:308–322.

65 Gajdosik CG. Ability of very young children to produce reliable isometric force measurements. *Pediatr Phys Ther*. 2005;17:251–257.

66 Hébert LJ, Maltais DB, Lepage C, Saulnier J, Crête M, Perron M. Isometric muscle strength in youth assessed by hand-held dynamometry: a feasibility, reliability, and validity study. *Pediatr Phys Ther*. 2011;23:289–299.

67 King-Dowling S, Proudfoot NA, Cairney J, Timmons BW. Validity of field assessments to predict peak muscle power in preschoolers. *Appl Physiol Nutr Metab*. 2017;42:850–854.

68 Latorre Román PÁ, Mora López D, Fernández Sánchez M, Sánchez JS, Coronas FM, García-Pinillos F. Test-retest reliability of field-based physical fitness assessment for children aged 3-6 years. *Nutr Hosp*. 2015;32:1683–1688.

69 Latorre-Román PÁ, García-Pinillos F, Mora-López D. Reference values of standing long jump in preschool children: a population-based study. *Pediatr Exerc Sci*. 2017;29:116–120.

70 Lucas BR, Latimer J, Doney R, et al. The Bruininks-Oseretsky Test of Motor Proficiency-Short Form is reliable in children living in remote Australian aboriginal communities. *BMC Pediatr*. 2013;13:135.

71 Macfarlane TS, Larson CA, Stiller C. Lower extremity muscle strength in 6- to 8-year-old children using hand-held dynamometry. *Pediatr Phys Ther*. 2008;20:128–136.

72 Marmis C, Montoye HJ, Cunningham DA, Kozar AJ. Reliability of the multi-trial items of the AAHPER youth fitness test. *Res Q*. 1969;40:240–245.

73 Molenaar HM, de Kraker M, Zuidam JM, Hovius SER, Stam HJ, Selles RW. Visual feedback and weight reduction of a grip strength dynamometer do not increase reliability in healthy children. *J Hand Ther*. 2010;23:272–279.

74 Morrow JR, Martin SB, Jackson AW. Reliability and validity of the FITNESSGRAM: quality of teacher-collected health-related fitness surveillance data. *Res Q Exerc Sport*. 2010;81:S24–S30.

75 O'Connell DG, O'Connell JK, Garrett ML, Adams N, Patterson B, Spencer E. Isometric strength and dynamic back extensor endurance are unrelated in children ages 6-10 years: a pilot study. *Percept Mot Skills*. 2004;99:1290–1294.

76 Ramírez-Vélez R, Rodrigues-Bezerra D, Correa-Bautista JE, Izquierdo M, Lobelo F. Reliability of health-related physical fitness tests among Colombian children and adolescents: the FUPRECOL study. *PLoS One*. 2015;10:e0140875.

77 Steenman K, Verschuren O, Rameckers E, Douma-van Riet D, Takken T. Extended reference values for the muscle power sprint test in 6-to 18-year-old children. *Pediatr Phys Ther*. 2016;28:78–84.

78 Vanhelst J, Béghin L, Fardy PS, Ulmer Z, Czaplicki G. Reliability of health-related physical fitness tests in adolescents: the MOVE Program. *Clin Physiol Funct Imaging*. 2016;36:106–111.

79 Yin L, Tang C, Tao X. Criterion-related validity of a simple muscle strength test to assess whole body muscle strength in Chinese children aged 10 to 12 years. *Biomed Res Int* 2018;Article ID 2802803:11.

80 Dekkers KJ, Rameckers EA, Smeets RJ, Janssen-Potten YJM. Upper extremity strength measurement for children with cerebral palsy: a systematic review of available instruments. *Phys Ther*. 2014;94:609–622.

81 Stark T, Walker B, Phillips JK, Fejer R, Beck R. Hand-held dynamometry correlation with the gold standard isokinetic dynamometry: a systematic review. *PMR*. 2011;3:472–479.

82 Castro-Piñero J, Artero EG, España-Romero V, et al. Criterion-related validity of field-based fitness tests in youth: a systematic review. *Br J Sports Med*. 2010;44:934–943.

83 Menz SM, Hatten K, Grant-Beuttler M. Strength training for a child with suspected developmental coordination disorder. *Pediatr Phys Ther*. 2013;25:214–223.

84 Hyde C, Wilson PH. Online motor control in children with developmental coordination disorder: chronometric analysis of double-step reaching performance. *Child Care Health Dev*. 2011;37:111–122.

85 Jucaite A, Fernell E, Forssberg H, Hadders-Algra M. Deficient coordination of associated postural adjustments during a lifting task in children with neurodevelopmental disorders. *Dev Med Child Neurol*. 2003;45:731–742.

86 Bieber E, Smits-Engelsman BC, Sgandurra G, et al. Manual function outcome measures in children with developmental coordination disorder (DCD): systematic review. *Res Dev Disabil*. 2016;55:114–131.

Downloaded from https://academic.oup.com/ptj/article/100/10/1825/5868019 by University of Groningen user on 16 October 2020

Appendix

Search Terms

Step 1

Keywords target population

DCD, developmental coordination disorder, motor development problems, Motor Skills Disorders, clumsy child syndrome, minimal brain dysfunction, clumsiness, clumsy, minimal brain dysfunction, minor neurological dysfunction, motor delay disorder, motor coordination difficulties, motor learning difficulties, motor coordination problems, mild motor problems, sensorimotor difficulties, sensory integrative dysfunction, physical awkwardness, physically awkward, psychomotor disorders, apraxia, developmental dyspraxia, perceptual motor dysfunction, perceptual-motor impairment, non-verbal learning disability, non-verbal learning disorder, coordination disorder, motor proficiency, low motor competence, motor impairment, motor difficulties

Keywords construct of interest

Anaerobic, anaerobic capacity, anaerobic endurance, endurance, strength, muscle strength, isometric strength, muscle force, muscle power, functional strength, physical fitness.

Step 2

Keywords target population

Child

Keywords construct of interest

All the measurements and tests from step 1

Measurement properties

(structural, construct, content, convergent, discriminative, criterion) validity, internal consistency, (interrater, intrarater, test-retest) reliability, responsive(ness), measurement error, sensitivity, specificity, psychometrics (properties)

Downloaded from https://academic.oup.com/ptj/article/100/10/1825/5868019 by University of Groningen user on 16 October 2020