Downs, SJ, Boddy, LM, McGrane, B, Rudd, JR, Melville, CA and Foweather, L

Motor competence assessments for children with intellectual disabilities and/or autism: a systematic review.

BMJ Open Sport & Exercise Medicine

Motor competence assessments for children with intellectual disabilities and/or autism: a systematic review

Samantha J Downs,[1] Lynne M Boddy,[1] Bronagh McGrane,[2] James R Rudd,[1] Craig A Melville,[3] Lawrence Foweather[1]

To cite: Downs SJ, Boddy LM, McGrane B, *et al*. Motor competence assessments for children with intellectual disabilities and/or autism: a systematic review. *BMJ Open Sport & Exercise Medicine* 2020;**0**:e000902. doi:10.1136/bmjsem-2020-000902

► Supplemental material is published online only. To view please visit the journal online (http://dx.doi.org/10.1136/bmjsem-2020-000902).

Received 24 July 2020
Revised 29 September 2020
Accepted 2 November 2020

Check for updates

[1]Research Institute for Sport and Exercise Sciences, Liverpool John Moores University, Liverpool, UK
[2]School of Arts Education and Movement, Dublin City University, Dublin, Ireland
[3]Institute of Health and Wellbeing, University of Glasgow, Glasgow, UK

Correspondence to
Lawrence Foweather;
L.Foweather@ljmu.ac.uk

ABSTRACT

Objective Gross motor competence is essential for daily life functioning and participation in physical activities. Prevalence of gross motor competence in children with intellectual disabilities (ID) and/or autism is unclear. This systematic review aimed to identify appropriate assessments for children with ID and/or autism.

Design & data sources An electronic literature search was conducted using the EBSCOhost platform searching MEDLINE, Education Research Complete, ERIC, CINAHL Plus and SPORTDiscus databases.

Eligibility criteria Included studies sampled children with ID and/or autism aged between 1 and 18 yrs, used field-based gross motor competence assessments, reported measurement properties, and were published in English. The utility of assessments were appraised for validity, reliability, responsiveness and feasibility.

Results The initial search produced 3182 results, with 291 full text articles screened. 13 articles including 10 assessments of motor competence were included in this systematic review. There was limited reporting across measurement properties, mostly for responsiveness and some aspects of validity. The Bruininks–Oseretsky Test of Motor Proficiency-2 followed by The Test of Gross Motor Development-2 demonstrated the greatest levels of evidence for validity and reliability. Feasibility results were varied, most instruments required little additional equipment (n=8) and were suitable for a school setting, but, additional training (n=7) was needed to score and interpret the results.

Conclusion This review found the BOT-2 followed by the TGMD-2 to be the most psychometrically appropriate motor competency assessments for children with ID and/or autism in field-based settings. Motor competence assessment research is limited for these cohorts and more research is needed.

PROSPERO registration number CRD42019129464.

Summary box

- ► Gross motor competence is an essential foundation for daily life functioning and to build more complex skills necessary for sport-specific activities or physical activity participation.
- ► It is unclear which motor competence tool is reliable, valid and feasible for use with children with intellectual disabilities and/or autistic spectrum disorder.
- ► This review identified the Bruininks-Oseretsky Test of Motor Proficiency, Second Edition, followed by the Test of Gross Motor Development-2 as assessments with the best population-specific measurement properties to date.
- ► While there are many assessments of gross motor competence, more research is needed on their validity, reliability and feasibility for use in children with ID and/or autism.

INTRODUCTION

Motor competence refers to performing goal-directed human movements in a coordinated, accurate and relatively error-free manner.[1–3] Fundamental movement skills (FMS; also termed foundational/gross motor skills) such as stability (eg, balancing), locomotor (eg, jumping) and object-control (eg, catching) skills are an important constituent of gross motor competence.[4 5] The development of gross motor competencies, including FMS, is considered an essential foundation for daily life functioning and to build more complex skills necessary for sport-specific activities or physical activity (PA) participation.[5–10] Gross motor competence promotes positive PA and health trajectories in children and adolescents,[2 11–15] including those with intellectual disabilities (ID) and Autistic Spectrum Condition (ASC).[16 17] Compared to typically developing peers, children with ID and ASC engage in less PA,[18 19] have low fitness levels[20] and greater rates of overweight and obesity.[19 21 22] Thus, motor competence deficits could further exacerbate these health inequalities. It is therefore imperative to monitor gross motor competence in children with ID and ASC to identify and diagnose motor development problems and support targeted interventions.

ID and ASC are neurodevelopmental disabilities. ID is characterised by impairments to intellectual and adaptive functioning,[23] which presents with difficulties in comprehending new and complex information, learning and applying cognitive, language, motor and social skills, as well as challenging behaviours.[24] ASC, on the other hand, is a permanent neurodevelopmental condition characterised by social, communication and interaction difficulties, and by repetitive and/or restrictive patterns of behaviour.[25 26] These social deficits observed in ASC can also present with more severe forms of ID, while individuals with ASC can display features that overlap with ID, such as taking longer to understand information.[27] This leads to diagnostic challenges for clinicians in distinguishing between ID as its own diagnoses, ID with an additional diagnosis of ASC, and ASC only, particularly in infants and very young children where some of these abilities are yet to emerge.[27] Furthermore, individuals with both ASC and ID appear to have a common genetic aetiology, with up to 50% of the children with ASC thought to have comorbid ID.[28] Therefore, given the similarities and complexities around the clinical manifestation of ASC and ID and the associated diagnostic challenges, the present systematic review focuses on children with ID and/or ASC.

Clinicians, physical therapists, physical educators and scholars require field-based assessments of gross motor competence that are valid, reliable and feasible to provide them with useful information for clinical, educational, and research purposes.[29 30] Validity, defined as 'the degree to which (an instrument) is an adequate reflection of the construct to be measured'[31] (eg, content validity). Reliability refers to 'the degree to which the measurement is free from measurement error'[31] (ie, test–retest reliability, intra- and inter-rater reliability). Feasibility refers to the usability of the assessment, including ease of administration, training or equipment requirements, cost and the length of time required.[32 33] Thus, the assessment must be acceptable to children and adolescents, researchers and/or professionals. Assessments should also be responsive and able to detect changes in gross motor competence, in order to monitor growth and development, and to evaluate the impact of interventions.[30] Information on these measurement properties is important as it influences the selection of the appropriate gross motor competence assessment for the intended purpose in the population of interest.

Several reviews have examined the measurement properties of gross motor assessment tools for use with typically developing children and adolescents.[34–37] While there is no 'gold standard' measure of gross motor competence, these reviews indicate the availability of an abundance of process- or product-oriented measures, or hybrid approaches. Process measures focus on the analysis of movement technique. This provides rich data on movement quality but extensive training is typically required due to the higher expertise needed for scoring skill criteria as present or absent. Product measures, which focus on the outcome of the movement (eg, running velocity, number of catches), are more objective, easy to score and less time consuming, and consequently have more limited training requirements (for a more detailed guide, see.[38 39]) While these reviews highlight a number of valid, reliable and feasible tools for use within typically developing children, it should not be assumed that these tools are appropriate for use with children with ID and/or ASC and more specific research is warranted.

Children with ID and/or ASC have complex needs, including communication issues, a limited attention span and ability to retain information. These populations may need to receive instructions and information in a different way to typically developing children,[40–42] thus requiring adapted forms of gross motor assessment administration. For instance, Wilson, Enticott and Rinehart[43] adapted the 3rd edition of the Test of Gross Motor Development (TGMD) to include visual support for those with ASC as it is known that children with ASC may have a preference for visual learning. They found the TGMD-3 raw scores of children with ASC were significantly lower than typically developing peers, however, their raw scores significantly improved using the TGMD-3 visual support protocol compared to the TGMD-3 traditional protocol. This indicates that children with ID and/or ASC may not understand the assessment requirements in existing assessments,[44] which could lead to the documentation of greater deficits in gross motor competencies in these populations relative to typically developing children than truly exists.

A number of studies have assessed gross motor competence in children with ID and/or ASC.[45–47] Despite a growth in research in these populations over the last decade, studies have used different assessment tools such as the TGMD-2 (eg,[48 49]) TGMD-3 (eg,[50]) or the Bruininks-Oseretsky Test of Motor Proficiency-2 (BOTMP-2: for example, [51]) which means that the results are not directly comparable and hinders broader interpretations of gross motor competence levels. It also highlights that there has been difficulty deciding on an assessment tool which may be most appropriate for use with children with ID and/or ASC, as these assessment tools were not originally designed for use with these populations. This is important as it is recommended that the quality of an assessment tool should be established in the target population in which the measure will be administered.[52] Furthermore, some of the available evidence used measures of only one dimension of FMS (typically locomotor or object-control), providing a narrow picture of gross motor competence, while the reliability of the assessments was unclear.[47] To overcome these methodological weaknesses in the literature, more clarity is needed regarding the measurement properties of gross motor competence assessment tools in children with ID and/or ASC.

The purpose of this systematic review was to evaluate the measurement properties of field-based assessments of gross motor competence for use in children with ID and/or ASC aged 3 to 18 years old. This information is needed to help professionals (educators/clinicians) and researchers determine the most appropriate and feasible tool for use with this specific population.

METHOD

The Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA)[53] guided the methodology and reporting of this study. The study protocol was registered with PROSPERO, registration no: CRD42019129464.

Inclusion criteria

Target population included:

- Children and adolescents with ID and/or ASC (mild, moderate, severe and profound).
- Participants aged between 3 and 18 years old.

Studies were included if they:

- Included a representative sample of the target population (>50% of sample had ID/ASC and >50% of sample <18 years old).
- Included a field-based assessment tool of gross motor competence (studies that included both fine and gross motor assessments were included if the data could be separated between fine and gross motor competence).
- Included measurement properties data of a gross motor competence tool.
- Were published in English language in a peer-reviewed journal.
- Were published between January 1937 and October 2019.

Exclusion criteria

During screening studies were excluded if:

- Included a laboratory assessment of gross motor competence.
- Included only fine motor skill assessments.
- Full text articles were not available.

Literature search and study selection

An electronic literature search was conducted using the EBSCOhost platform searching MEDLINE, Education Research Complete, ERIC, CINAHL Plus and SPORTDiscus databases. These databases were searched independently and included publications from January 1937—October 2019. The search strategies applied in the databases included combinations of key search terms from four subheadings; population (Child* OR adolesc* OR youth* OR teenage* OR girl* OR boy* OR preschool* OR juvenile OR p$ediatric) AND disability (Disab* OR autis* OR 'down syndrome' OR (intellectual* OR learning OR develop* OR mental* OR special OR additional OR cognitive) NEAR/2 (disab* OR disorder* OR impair* OR difficult* OR handicap* OR deficien* OR subnorm* OR delay* OR retard* OR needs)) AND motor competence (Coordination OR agil* OR balanc* OR ('fundamental movement' OR physical* OR movement OR motor OR locomotor OR 'object control' OR 'gross motor' OR stability OR actual) NEAR/2 (skill* OR compet* OR develop* OR abil* OR proficien* OR learning OR assess*)) AND assessment tool (Assess* OR measure* OR test* OR tool OR battery OR instrument* OR evaluat* OR valid* OR reliab* OR feasib* OR responsiveness OR 'psychometric test' OR 'measurement properties'). Boolean searches were also used.

Once the initial literature search was completed, the lead author removed all duplicates and remaining papers were transferred to Covidence (online systematic review software) for screening. Title and abstract screening of all articles was conducted by the lead author, a second researcher (last author) also independently screened 20% of the article titles and abstracts. Disagreements were resolved through discussing with a third researcher (second author). Full-texts were retrieved, screened and labelled 'yes', 'no' or 'maybe' for inclusion by the first author. The last author checked 10% of the articles labelled 'no' and all of the articles labelled 'yes'. Articles labelled 'maybe' were discussed between the first and last author until a consensus on the decision was made, with the involvement of the second author where necessary. In addition to the electronic literature search, researchers also checked reference lists of included papers and searched author bibliographies.

Data extraction

The first author independently extracted individual study data relating to: study information (title, authors, year and country of publish, study design and environment), participant information (sample size, age, sex, disability diagnosis and severity, inclusion/exclusion criteria, body mass index [BMI] and weight status), assessment purpose and administration, measurement properties of tools (reliability, validity, responsiveness and feasibility). Data extraction was checked for accuracy by *co-authors* and inconsistencies were resolved with discussion between the first and last author.

Quality appraisal

A quality appraisal tool for rating the measurement properties of motor competence assessments was developed on the basis of previous checklists (see online supplemental file 1).[30 31 52 54] In addition, to examine the feasibility of assessments, a utility matrix was developed using criteria gathered from previous recent systematic reviews exploring related concepts (see online supplemental file 2).[32 33] Risk of bias across studies was assessed using a modified version of the Grading of Recommendations Assessment, Development and Evaluation (GRADE) approach.[52 55]

RESULTS

Figure 1 provides the PRISMA flow diagram of the search and screening process. 13 articles were identified for inclusion including 10 unique instruments: the Test of Gross Motor Development 2nd Edition (TGMD-2),[56] Test of Gross Motor Development 3rd Edition (TGMD-3) [traditional and visual],[56] Bruininks-Oseretsky Test of Motor Proficiency Second Edition (BOT-2),[57] Movement Assessment Battery for Children 2nd Edition (MABC-2),[58] Peabody Developmental Motor Scales 2nd Edition

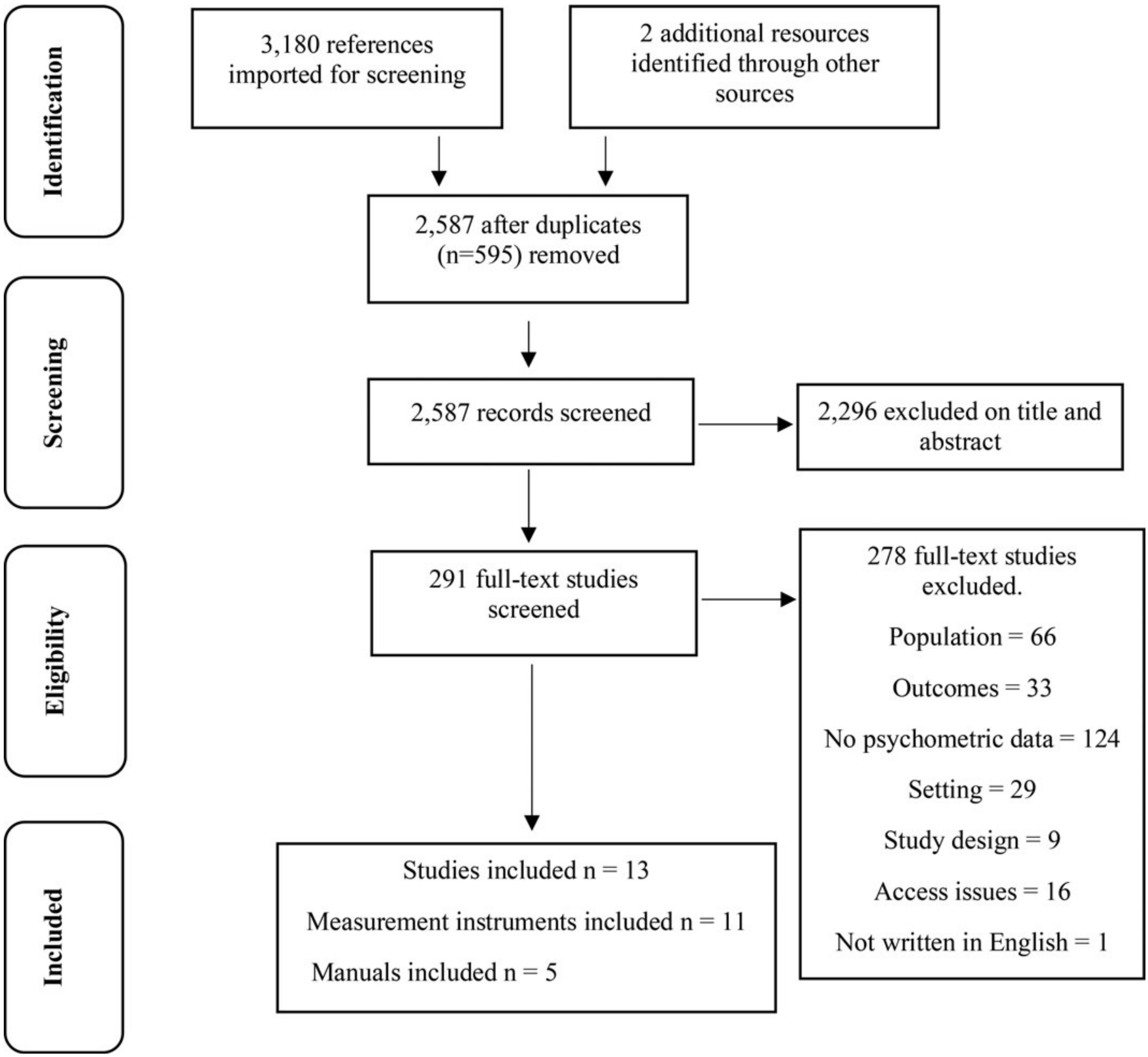

Figure 1 PRISMA flow diagram showing identification and screening process.

(PDMS-2),[59] Test of Motor Proficiency,[60] Cratty Six-Category Gross Motor Test,[61] Data-Based Dance Skills Placement Test,[62] Ages and Stages Questionnaire-2[63] and The Four Square Step Test.[64] Five corresponding manuals were identified and added to the final pool for the TGMD-2, TGMD-3, BOT-2, MABC-2 and the PDMS-2 tools; the remaining tools either did not have a manual or they could not be located.

Tool characteristics

Table 1 shows the characteristics of the 10 motor competence assessment tools documented in the included studies and manuals. Five tools are suitable for use with children (TGMD-2,[56] TGMD-3,[56] PDMS-2,[59] Ages and Stages Questionnaire-2,[63]), three tools with children and adolescents (BOT-2,[57] MABC-2,[58] Test of Motor Proficiency,[60]), with the remainder of tools suitable for youth and adult populations (Cratty Six Category Gross,[61] Dance Skill Placement Test,[62] Four Square Step Test.[64]). With the exception of the TGMD-2, TGMD-3 traditional and visual, the Four Square Step Test, Data-Based Dance Skill Placement and the Cratty Six Category Gross, all other tools measure at least one other construct in addition to gross motor competence such as fine motor skills or strength. Despite differences in test administration and structure, the tools tend to include similar items within the gross motor skill subsets, specifically object control skills such as, catching, throwing and kicking, or locomotor skills such as walking, running and jumping. Further similarities can be seen in strength tasks for the BOT-2

Table 1 Characteristics of the gross motor competence assessments

Instrument	Constructs	Subscales tested (items)	Age band	Scoring
Test of Gross Motor Development-2 (TGMD-2)[56]	Gross motor functioning	Locomotor (6) Object Control (6)	3–10 yrs	*Process* Every skill is evaluated on three to five performance criteria, 2 trials are summed for each skill. 0= if criteria are not preformed 1= if criteria are performed correctly
Test of Gross Motor Development-3 (TGMD-3)[56] [traditional & visual]	Gross motor skill performance	Locomotor (6) Ball skills (7)	3–10 yrs	*Process* Every skill is evaluated on three to five performance criteria, 2 trials are summed for each skill. 0= if criteria are not preformed 1= if criteria are performed correctly
Bruininks-Oseretsky Test of Motor Proficiency, Second Edition (BOT-2)[57]	Motor skills	Fine Motor Precision (7) Fine Motor Integration (8) Manual Dexterity (5) Bilateral Coordination (7) Balance (9) Running Speed & Agility (5) Upper-Limb Coordination (7) Strength (5)	4–21 yrs	*Product* Motor skills are quantified based on the results of goal-directed activities. A raw score for item outcome may be a drawing, a number of correct activities performed, a number of seconds to complete a task, and/or a complete/incomplete task. A scoring form is used to convert raw scores into point scores.
Movement Assessment Battery for Children—2nd Edition (Movement ABC-2)[58]	Motor skills	Skills differ for each age band Manual Dexterity (3) Aiming & Catching (2) Balance Static (1) Balance Dynamic (2)	3–6 yrs 7–10 yrs 11–16 yrs	*Product* Item performance may be a number of points, a number of performance correct or number of errors performed, and number of seconds to complete task.
Peabody Developmental Motor Scales, Second Edition (PDMS-2)[59]	Motor skills	Reflexes (8) Stationary (30) Locomotion (89) Object Manipulation (24) Grasping (26) Visual-Motor Integration (72)	0–5 yrs	*Product* 3 point scoring system per item performance 2) The child performs the item according to the criteria specified for mastery 1) The child's performance shows a clear resemblance to the item mastery criteria but does not fully meet the criteria 0) The child cannot or will not attempt the item, or the attempt does not show that the skill is emerging
Test of Motor Proficiency[60]	Motor skills	Visual-motor control (6) Upper limb speed & dexterity (5) Running speed & agility (1) Balance (5) Bilateral coordination (4) Strength (2) Upper limb coordination (15)	6–17 yrs	*Product* For each item the child's performance is assigned a numerical score (point score). Scoring for each item is different due to variations in performing each item under different subscales. Points for scoring individual items vary from 0–4 to 0–8 within as well as across the subscales.

Continued

Table 1 Continued

Instrument	Constructs	Subscales tested (items)	Age band	Scoring
Cratty Six-Category Gross Motor Test[61]	Gross motor skills	Body perception Gross agility Balance Locomotor agility Ball tracking Ball handling	4–24 yrs	*Product* Two six-item tests, one at the level of obvious motor problems and another at the level of mild motor problems. Scores (0–5) for each test item, maximum 10 points per category, and 60 points in total. All tasks in level one are administered and if his/her average score at this level is 4.0 or better, the tester administers the tasks in level two. If any single task in level one reached a score of 5.0, the participant is given the task in the corresponding category in level two.
Data Based Dance Skills Placement Test[62]	Gross motor skills	Locomotor and axial/ gymnastics (dance relevant) skills	Not reported	*Product* Scoring of the Data Based Dance Skills Placement Test consists of counting the number of phases and steps correctly performed and assigning 1 point to each. The items are scored on a Likert-scale (ranging from 0 to 4) including: always, little difficulty, some difficulty, great difficulty, and never. The ratings corresponded with numerical values and objective criteria. A rating of always indicated the student was able to perform the skill all four times or that a balance was held for four seconds.
Ages and Stages Questionnaire-2[63]	Motor skills (Screen developmental performance)	Communication skills Gross motor skills Fine motor skills Problem-solving skills Personal social skills	4–60 months	*Product* Parents answer questions surrounding whether children can complete tasks (yes =10 points, sometimes =5 points,no =0 points), minimum score of 0 and maximum score of 60 per domain. Sum scores compared to norm data.
The Four Square Step Test[64]	Gross motor (Dynamic standing balance)	Four Square Step Test (1)	+65 yrs	*Product* The score is recorded as the time taken to complete the item sequence. Two trials are completed with the best time taken as the score.

and Test of Motor Proficiency tools, as well as dexterity tasks for the BOT-2, PDMS-2 and Test of Motor proficiency. Five tools (BOT-2; PDMS-2; Test of Motor Proficiency; Cratty Six Category Gross; Ages and staged Questionnaire-2) assess all three gross motor skill domains: stability, object-control and locomotor skills. It should be noted that the Ages and Stages Questionnaire-2 is the only tool within this study that is a questionnaire and covers both fine and gross motor skills. All tools use product-oriented scoring procedures, with the exception of the process measures of TGMD-2 and TGMD-3.

Study characteristics

The characteristics of included studies in the review are described in table 2. The studies were conducted in four continents; Asia (n=7), Australia (n=1), Europe (n=3) and USA (n=2). Typically, data collection took place indoors within school grounds in an open area such as a multi-use hall or sports hall (n=10), working with children individually or in small groups (2–4 children). Study participants included individuals with Down syndrome (DS) (n=3 studies), ASC (n=2 studies) and ID (n=8 studies), with participant ages ranging from 1 year 10 months to 18-years old and study sample sizes ranging from 13 to 446 participants (average = 93 participants).

Validity, reliability and responsiveness

Online supplemental file 3 (validity and responsiveness) and online supplemental file 4 (reliability) displays the data extracted from the included studies with regards to the measurement properties of the assessments in children with ID and/or ASC. Table 3 presents a summary of the overall evidence for each of the 12 measurement properties, synthesising the outcomes of all studies for each measurement instrument. In general evidence was limited with few studies reporting across the range of measurement properties. Test-retest reliability, inter-rater reliability and internal consistency were most frequently evaluated, while no studies assessed content validity.

Validity data were available for nine of the ten assessments within 11 studies. Internal consistency was reported for seven tools (TGMD-2, TGMD-3 traditional, TGMD-3 visual, BOT-2, MABC-2, PDMS-2 and Test of Motor Proficiency) within 69% of studies. Discriminant validity (BOT-2), construct/structural validity (TGMD-2) and cross-cultural validity/measurement invariance (BOT-2) were reported for one tool within one study each; hypothesis testing for construct validity for three tools (TGMD-2 and TGMD-3 Traditional & Visual) within 15% of studies, and criterion/concurrent validity for seven tools (TGMD-3 Visual, BOT-2, MABC-2, PDMS-2, Test of Motor Proficiency, Ages and Stages Questionnaire-2 and the Four Square Step Test) within 38% of studies. Positive ratings were reported for all validity measurement properties with the exception of criterion/concurrent validity, which showed some indeterminate ratings for the PDMS-2 and the Ages and Stages Questionnaire-2. No tools had sufficient validity data to present a positive rating for all aspects of validity.

Reliability data were available for nine of the 10 identified measurement instruments. All studies that assessed elements of reliability were given a positive rating. Test-retest reliability was reported for 10 tools within 77% of studies (no test-retest data for Ages and Stages Questionnaire-2), intra-rater reliability for the TGMD-3 only (traditional and visual) within 8% of studies, and inter-rater reliability for eight tools (TGMD-2, TGMD-3 traditional, TGMD-3 visual, BOT-2, MABC-2, PDMS-2, Test of Motor Proficiency and the Four Square Step Test) within 62% of studies. There were no reliability data for the Ages and Stages Questionnaire-2 in this population.

Overall responsiveness results were mixed (table 3). Studies included in this review provided responsiveness results (including floor & ceiling effects) for three of the 10 measurement instruments reported within 23% of studies. The minimum important change and smallest detectable change results showed positive ratings while the AUC results were negative for BOT-2, MABC-2 and PDMS-2, this suggests that these tools are capable of detecting some change at an acceptable level but not a good level in this population. Floor and ceiling effects for the MABC-2 and PDMS-2 were both negative, while the BOT-2 reported a negative rating in one study[72] and a positive rating in another.[71] No other floor and ceiling effects were reported for other instruments.

Feasibility

The detailed information concerning the utility of each assessment is provided in online supplemental file 5. Table 4 presents a summary of the feasibility ratings for each assessment. Data are missing for the Ages and Stages Questionnaire—2, as this detail was not presented in the included studies and no manual could be located online or via the library, while the authors did not respond to contact made by the lead author. Results showed that only one measurement instrument (PDMS-2) took a long time to conduct (>60 mins), while the majority of instruments were rated positively, taking a short (<15 mins: Four Square Step Test; <30 mins: TGMD-2, TGMD-3, MABC-2) to moderate (30–60 mins: BOT-2, Test of Motor Proficiency, Data-based Dance Skill Placement Test) length of time to complete. The amount of space required to conduct different tests was more varied, with the TGMD-2 and TGMD-3 requiring the largest space (60 feet and 18.3 metres of clear space is required, respectively, for the run task alone). With regards to the equipment needed to administer the tests, over half of the instruments received positive ratings, with either the equipment required likely to be present in a typical school (Test of Motor Proficiency, Four Square Step Test) or minimal additional equipment required (TGMD-2, TGMD-3, BOT-2, PDMS-2, Cratty Six-Category Gross Motor Test). The Data-Based Dance Skill Placement Test requires no equipment. A large proportion of the instruments (60%) do not report specific qualifications required to carry out the assessments, however, the MABC-2 and PDMS-2 require a researcher with specific qualifications. Further, results imply that as well as administrators familiarising themselves with the examiner manuals, they would need further additional training for general administration and scoring purposes (0.5 day—1.5 days) (50%—TGMD-2, TGMD-3, BOT-2, MABC-2, Test of Motor Proficiency).

Risk of bias across studies

Online supplemental file 6 shows the risk of bias across studies for each measurement property per assessment. The quality of the evidence was low or very low across measurement properties for the Test of Motor Proficiency, Cratty Six-Category Motor Test, the Data-Based Dance Skill Placement Test, Ages & Stages Questionnaire-2 and the Four Square Step Test. The quality of the evidence for the remaining assessments (TGMD-2, TGMD-3, BOT-2, MABC-2, PDMS-2) was mixed and varied by measurement property, with moderate to high-quality evidence typically available for test-retest and inter-rater reliability, as well as internal consistency. The measurement properties were generally downgraded due to lack of available studies including a lack of high-quality studies, imprecision due to the total sample size included in the studies being below 100, and/or inconsistency in the similarity of results across studies for the measurement property.

Table 2 Characteristics of the studies reporting gross motor competence assessment measurement properties

Assessment	Study Authors & Country	Test Administration Setting	Study Population Information: Diagnosis	N & Sex	Age range (mean± SD)
Test of Gross Motor Development-2 (TGMD-2)[56]	Capio, Mak, Tse & Masters (Hong Kong)[65]	Premises of the advocacy group (DS) and a local community centre (TD)	DS	40 children (DS: n=20, 40% female; TD: n=20, 20% female)	DS: 6–11 years (7.1±2.9) TD: 6–11 years (7.3±2.5)
	Eguia, Capio & Simons (Philippines)[17]	On school premises	Mild-moderate ID	60 children, 15% female	5–14 years (9.6±3.2)
	Schott & Holfelder, (Germany)[66]	School gym	DS	36 children (DS: n=18, 39% female; TD n=18, 39% female)	7–11 years (9.0±0.9)
	Simons, Daly, Theodorou, Caron, Simons & Andoniadou (Belgium)[67]	During regular PE lessons—indoor facilities with PE teacher present	Cognitive delay, mild ID	99 children, 32% female	7–11 years (8.8±1.8) Boys, n=67 (9.7±1.2) Girls, n=32 (8.7±0.8)
Test of Gross Motor Development-3 (TGMD-3)[56] [Traditional]	Allen, Bredero, Damme, Ulrich and Simons (Australia)[68]	Indoor gymnasium at the School of Human Movement Studies, UQ	ASD	35 children (ASD n=14, 29% female; TD n=21, 43% female [9 were siblings of ASD participantss)	ASD: 4–10 years (7.4±2.0) TD: 4–10 years (7.3±1.8)
Test of Gross Motor Development-3 (TGMD-3)[56] [Visual]	Allen, Bredero, Damme, Ulrich and Simons (Australia)[68]	Indoor gymnasium at the School of Human Movement Studies, UQ	ASD	14 children, 29% female	4–10 years (7.8±1.9)
Bruininks-Oseretsky Test of Motor Proficiency, Second Edition (BOT-2)[57]	Pitetti, Miller & Loovis (USA)[69]	Classroom or gymnasium	Mild-moderate ID (IQ scores ranged from 45–70), no participants with DS.	123 children, 0 female	8–18 years (not reported)
	Wuang & Su (Taiwan)[70]	Participants school, home or hospital	Mild ID (n=64) moderate ID (n=36)	100 children, 41% female	4–12 years (6.9±2.1)
	Wuang, Lin & Su (Taiwan)[71]	Participants school, home, or paediatric occupational therapy unit	Mild ID (71.7%; IQ =55–70) moderate-severe ID (28.3%; IQ =25–54)	446 children, 40% female	4–18 years (9.4±4.0)
	Wuang, Su & Huang (Taiwan)[72]	As above	mild ID (n=101) moderate-severe ID (n=40)	141 children, 37% female	3–6 years (3.6±1.2)
Movement Assessment Battery for Children—2nd Edition (Movement ABC-2)[58]	Schott & Holfelder, (Germany)[66]	School gyms	DS	36 children (DS n=18, 39% female; TD n=18, 39% female)	7–11 years (9.0±0.9)

Continued

Table 2 Continued

Assessment	Study Authors & Country	Test Administration Setting	Study Population Information: Diagnosis	N & Sex	Age range (mean± SD)
	Wuang, Su & Huang (Taiwan)[72]	Participants school, home or paediatric occupational therapy unit	Mild ID (n=101) moderate-severe ID (n=40)	141 children, 37% female	3–6 years (3.6±1.2)
Peabody Developmental Motor Scales, Second Edition (PDMS-2)[59]	Vanvuchelen, Van Schuerbeeck & Braeken (Belgium)[73]	Not reported	ASD, IQ>70: none of the participants have ID	43 children, 23% female	22mths–4.5 years (3.3±0.7)
	Wuang, Su & Huang (Taiwan)[72]	Participants school, home or paediatric occupational therapy unit	Mild ID (n=101) moderate-severe ID (n=40)	141 children, 37% female	3–6 years (3.6±1.2)
Test of Motor Proficiency[60]	Kalgotra & Warwal (India)[60]	Special Schools but not specifically stated	Mild ID (n=26; IQ ranges from 50–65) Moderate ID (n=24; IQ ranges from 35 to 46)	50 children, sex not reported	6–17 years Mild ID (13.2±3.1) Moderate ID (11.7 ±2.6)
Cratty Six-Category Gross Motor Test[61]	Roswal, Sherrill & Roswal (USA)	Schools—SEN classes	Moderate ID, specifics not reported	35 children, 34% female. Data based dance group n=18, 33% female Creative dance group n=17, 35% females	11–16 years Data group (12.9±6.1) Creative group (13.5 ±6.1)
Data Based Dance Skills Placement Test[62]	Roswal, Sherrill & Roswal (USA)[74]	Schools—SEN classes	Moderate ID, specifics not reported	35 children, 34% female. Data based dance group n=18, 33% female Creative dance group n=17, 35% females	11–16 years Data group (12.9±6.1) Creative group (13.5 ±6.1)
Ages and Stages Questionnaire-2[63]	Vanvuchelen, Van Schuerbeeck & Braeken (Belgium)[73]	Not reported	ASD, IQ>70: none of the participants have ID	43 children, 23% female	22mths–4.5 years (3.3±0.7)
The Four Square Step Test[64]	Verma, Samuel & Aranha (India)[75]	Not reported	DS	13 children, 38% female	8–17 years (12.8±2.9)

ASD, Autistic spectrum disorder; DS, Down syndrome; ID, Intellectual disability; IQ, IQ; SEN, Special educational needs; TD, Typically developing.

DISCUSSION

This systematic review aimed to identify and evaluate the measurement properties of field-based assessments of gross motor competence for use in children with ID and/or ASC aged 3 to 18 years old. In general, there was a lack of published studies (n=13) concerning gross motor competence assessments in ID and/or ASC populations. Nevertheless, the results suggest assessments exist which are psychometrically sound to examine motor competence in these populations within a field-based setting (eg, a school). However, the quality of the evidence was limited by the small number of studies and low pooled sample size of the included studies for each measurement property per assessment. Only three tools were assessed across each measurement dimension (validity, reliability and responsiveness): BOT-2,[70 72] MABC-2[72] and PDMS-2,[72] while none of the studies included in this review reported across all 12 measurement properties. Two studies reported on six measurement properties covering four tools: the BOT-2, MABC-2, PDMS-2[72] and the

Table 3 Summary of appraisal of measurement properties of gross motor competence assessments

	Reliability			Validity							Responsiveness	
Assessment	Test-retest	Intra-rater	Inter-rater	Internal Consistency	Content validity	Discriminant validity	Construct/ structural validity	Hypothesis testing	Cross-cultural validity/ invariance	Criterion/ concurrent validity	Responsiveness	Floor & ceiling effects
TGMD-2	++	NR	+++	++++	NR	NR	+	+	NR	NR	NR	NR
TGMD-3	+	+	+	+	NR	NR	NR	+	NR	+	NR	NR
BOT-2	+++	NR	+	++	NR	+	NR	NR	+	+	+ +	+ -
MABC-2	+	NR	+	++	NR	NR	NR	NR	NR	+	+	–
PDMS-2	++	NR	++	++	NR	NR	NR	NR	NR	?+	+	–
Test of Motor Proficiency	+	NR	+	+	NR	NR	NR	NR	NR	+	NR	NR
Cratty Six-Category Gross Motor Test	+	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR
Data Based Dance Skill Placement Test	+	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR
Ages & Stages Questionnaire —2	NR	NR	NR	NR	NR	NR	NR	NR	NR	?	NR	NR
Four Square Step Test	+	NR	+	NR	NR	NR	NR	NR	NR	+	NR	NR

NR, not reported; +, positive rating; -, negative rating; ? = indeterminate rating. Number of symbols in each cell relates to the number of studies reporting the measurement property (eg, + =1 study; +++ =3 studies). TGMD-2= Test of Gross Motor Development 2nd Edition. TGMD-3= Test of Gross Motor Development 3rd Edition. BOT-2= Bruininks-Oseretsky Test of Motor Proficiency, Second Edition. MABC-2= Movement Assessment Battery for Children 2nd Edition. PDMS-2= Peabody Developmental Motor Scales 2nd Edition.

Table 4 Feasibility scores for gross motor competence assessments

Assessment tool	Time	Space	Equipment	Qualification	Training
TGMD-2	***	**	***	NR	**
TGMD-3	***	*	***	NR	**
BOT-2	**	**	***	NR	**
MABC-2	***	***	**	*	**
PDMS-2	*	***	***	*	NR
Test of Motor Proficiency	**	*	*	***	**
Cratty Six-Category Gross Motor Test	NR	**	***	NR	NR
Data Based Dance Skill Placement Test	**	***	None required	**	NR
Ages and Stages Questionnaire—2	NR	NR	NR	NR	NR
The Four Square Step Test	****	****	****	NR	NR

****Excellent, ***Good, **Fair, *Poor. NR, Not Reported.

TGMD-3 (traditional and visual).[68] However, most studies (71%) reported on three or less measurement properties, with test-retest reliability, inter-rater reliability and internal consistency properties most commonly reported. Taken together, the review has highlighted the lack of research conducted around the measurement of motor competence in children with ID and/or ASC and more research is warranted.

Of the 13 published studies in the review, most measurement instruments only appear in one (64%) or two (18%) papers. The BOT-2 and TGMD-2 both appeared within four studies and were therefore most commonly reported. These findings are consistent with systematic reviews of gross motor competence assessments for typically developing children.[33 35 37] These previous reviews include over 30 different gross motor competence assessments that, due to a lack of published data on measurement properties in children with ID or ASC, were not included in the current review. These include traditional style assessments such as CHAMPS Motor Skill Protocol,[76] and obstacle/circuit-based instruments such as Dragon Challenge[77] and MOBAK-5-6,[78] as well as the use of advanced technologies such as inertia sensors to measure competence.[79] Further research examining the validity, reliability and feasibility of these tools and new technologies in children with ID or ASC is needed to ascertain the suitability of these tools for use with this population.

Validity

Internal consistency was the most commonly reported form of validity (reported in 69% of studies), followed by criterion/concurrent validity (reported in 38% of studies). The TGMD-2 reported good internal consistency when used on populations with ID.[17 65–67] The BOT-2, MABC-2 and PDMS-2 also reported good internal consistency when used with ID and ASC populations.[66 70 72 73] Positive criterion/concurrent validity was observed for TGMD-3 visual when assessed against TGMD-3 traditional[68]; The Four Square Step Test when assessed against the Functional Reach Test[75]; the Test of Motor Proficiency when assessed against BASIC-MR Behavioural Assessment Scale For Indian Children With Mental Retardation, Part- A[60]; the BOT-2, MABC-2 and PDMS-2 when compared with one another.[72] In one study[73] criterion/concurrent validity data were unclear resulting in indeterminate ratings for the Ages and Stages Questionnaire-2 when assessed against the PDMS-2.

Many other aspects of validity were underreported: content validity was not reported for any tools, while discriminant validity (BOT-2,[71]) construct/structural validity (TGMD-2[67]) and cross-cultural validity/measurement invariance (BOT-2[71]) properties were rarely reported. These findings differ somewhat from similar systematic reviews conducted in TD populations.[36 37] These reviews found that construct validity was the most commonly assessed aspect of validity and that generally assessments had sound quantitative evidence for proposed factor structures for motor constructs. Our finding that content validity was reported the least is in agreement with Hulteen *et al*,[37] yet in contrast to Scheuer *et al*,[36] who found that content validity was the second most commonly reported form of validity (60% of studies). It should be noted that the differences observed between these reviews and ours may be linked to methodological factors. Regardless, the lack of information concerning content validity and the lack of testing for validity within studies included in the current review may suggest that it is assumed that tests developed for use with TD children will be valid and appropriate for use children with ID and ASC. Given the importance of using assessment tools that are validated for use with that specific population,[52] more validity testing is required in children with ID or ASC to make definitive statements regarding the validity of assessments.

Reliability

Test-retest reliability was the most commonly reported property for reliability, followed by inter-rater reliability, while only a single study examined intra-rater reliability.[68] As noted by Hulteen *et al*,[37] it is interesting that test-retest reliability was reported the most given that it is more time consuming, with greater burden for the participant, as this construct requires data collection on at least two time points for each participant, ideally 2 weeks apart. In comparison inter-rater and intra-rater reliability constructs can be checked during the same testing session. Of the 10 measurement instruments reviewed in the current study, the TGMD-2, BOT-2 and PDMS-2 reported the highest levels of reliability. These studies included researchers,[66 70 72] licenced physiotherapists,[17] occupational therapists,[70 72] university students,[69] physical education specialists and psychomotor therapists[67] when testing reliability. These three tools consistently demonstrated strong inter-rater[17 66 67 72 73] and test-retest reliability (ICC or r>0.70).[17 67 69 70 72 73] The TGMD-3 was the only tool that had data for all three reliability measurement properties,[68] while the Ages and Stages Questionnaire—2 did not have any reliability data.[73]

Responsiveness

Assessments should be responsive and able to detect changes in gross motor competence, in order to monitor growth and development, and to evaluate the impact of interventions.[30] In the current review, responsiveness and floor and ceiling effects were reported for three (BOT-2, MABC-2 and PDMS-2) of the 10 measurement instruments. Responsiveness properties were reported in two papers for the BOT-2 and demonstrated positive ratings in both studies, while the floor & ceiling effects showed mixed results (+ −), suggesting the tool can detect change but only at an acceptable level. Similar to previous systematic reviews in TD children,[34–37] in general the reporting of responsiveness measurement properties within studies included in this review were limited. This could be due to the low prevalence of reporting.[37] Alternatively, testing for responsiveness requires researchers to conduct longitudinal or experimental studies, which are more difficult and time consuming to conduct. Testing for responsiveness has been identified as an important area for future research.[35]

Feasibility

Given the wide range of disabilities within SEN school settings, ensuring that field-based assessments are feasible for researchers and professionals to use in school settings is important. Different factors can influence the feasibility of assessments,[33] such as time, space, environment and administrator expertise/training.[34] As the development of gross motor competence is optimal during childhood,[6] the feasibility was assessed based on primary school settings and the majority of included studies (71%) were conducted within this environment. Feasibility varied from one instrument to the next as well as within the instruments. For example, the PDMS-2 scored poorly for time and qualifications required to complete the test, yet the amount of space and equipment needed scored well. Instruments that scored poorly with regards to 'time' tended to assess a wider array of gross motor domains and skills (BOT-2, PDMS-2 and Test of Motor Proficiency). When conducting these more detailed assessments, administrators could break testing sessions up by subsets to alleviate some of the time burden for participants.[80] As feasibility information was incomplete for eight of 10 measurement instruments, it is not possible to make a fair and 'final' judgement. However, limited results suggest the Test of Motor Proficiency[60] scored the highest and the PDMS-2[59] scored the lowest for feasibility, with the need for administrator training unclear. The Test of Motor Proficiency was specifically developed for the assessment of gross and fine motor skills of children with mild-moderate ID within SEN schools, perhaps explaining why it is the most feasible instrument within this study.[60] More research is required to establish the feasibility of existing assessments for use with children with ID and ASC in primary school settings.

Strengths and limitations

Strengths of this paper include the adherence to PRISMA[53] guidelines and the appraisal of a wide range of measurement properties to examine the quality of the assessments in terms of validity, reliability, responsiveness and feasibility. Limitations of the review include the exclusion of fine motor skills, which are an important constituent of motor development, and laboratory or clinical-based assessments, as we wanted to focus on gross motor competence assessments that can be administered by researchers and professionals in the children's natural environments such as home, school and community settings. Further, only papers written in English within peer-reviewed journals were included, meaning we may have missed some relevant assessments.

CONCLUSION

This is the first paper to systematically review the validity, reliability, responsiveness and feasibility of assessments of gross motor competence in children with ID and/or ASC. While 10 instruments were identified, the available evidence was of mixed quality: literature was sparse with many measurement properties unreported or not yet examined in the target population. The limited evidence available suggests that the BOT-2,[57] followed by the TGMD-2,[56] have the strongest measurement properties to support use of these assessments with children with ID and/or ASC to date. Assessments developed specifically for use with children with ID and/or ASC such as the Test of Motor Proficiency[60] scored highest for feasibility, supporting the importance of using assessment tools designed for use with this specific population. More population-specific research is required to establish the validity, reliability, responsiveness and feasibility of existing assessments for use with children with ID and ASC in primary school settings.

Contributors SD and LF conceived the study and secured the funding. All authors were involved in the study design and set the scope of the review, agreed articles for inclusion, and made contributions to the quality appraisal of assessments. SD and LF wrote the manuscript. All other *coauthors* read, reviewed, commented and approved the final draft.

Funding This study was funded by the Baily Thomas Charitable Fund (award ref: TRUST/VC/AC/SG/4771-7962).

Competing interests None declared.

Patient consent for publication Not required.

Ethics approval Not required.

Provenance and peer review Not commissioned; externally peer reviewe.

Data availability statement All data relevant to the study are included in the article or uploaded as supplemental information.

ORCID iDs

Lynne M Boddy http://orcid.org/0000-0002-7477-4389
Bronagh McGrane http://orcid.org/0000-0001-6891-7729
James R Rudd http://orcid.org/0000-0003-1546-576X
Craig A Melville http://orcid.org/0000-0001-7234-2382
Lawrence Foweather http://orcid.org/0000-0001-9851-5421

REFERENCES

1 Anson G, Elliott D, Davids K. Information processing and constraints-based views of skill acquisition: divergent or complementary? *Motor Control* 2005;9:217–41.
2 Robinson LE, Stodden DF, Barnett LM, *et al.* Motor competence and its effect on positive developmental trajectories of health. *Sports Med* 2015;45:1273–84.
3 Rudd JR, Barnett LM, Butson ML, *et al.* Fundamental movement skills are more than run, throw and catch: the role of stability skills. *PLoS One* 2015;10:e0140224.
4 Logan SW, Ross SM, Chee K, *et al.* Fundamental motor skills: a systematic review of terminology. *J Sports Sci* 2018;36:781–96.
5 Gallahue DL, Ozmun JC, Goodway J. *Understanding motor development: Infants, children, adolescents, adults*. 7th edn. New York, NY: McGraw-Hill Education, 2012.
6 Clark JE, Metcalfe JS. The mountain of motor development: a metaphor. In: Clark JE, Humphrey J, eds. *Motor development: research and reviews*. Reston, VAA: NASPE Publications, 2002: 163–90.
7 De Meester A, Stodden D, Goodway J, *et al.* Identifying a motor proficiency barrier for meeting physical activity guidelines in children. *J Sci Med Sport* 2018;21:58–62.
8 Seefeldt V. Developmental motor patterns: implications for elementary school physical education. In: Nadeau CH, Halliwell WR, Newell K, eds. *Psychology of motor behavior and sport*. Champaign, IL: Human Kinetics, 1979: 314–23.
9 Stodden DF, Goodway JD, Langendorfer SJ, *et al.* A developmental perspective on the role of motor skill competence in physical activity: an emergent relationship. *Quest* 2008;60:290–306.
10 Hulteen RM, Morgan PJ, Barnett LM, *et al.* Development of foundational movement skills: a conceptual model for physical activity across the lifespan. *Sports Med* 2018;48:1533–40.
11 Bremer E, Cairney J. Fundamental movement skills and health-related outcomes: a narrative review of longitudinal and intervention studies targeting typically developing children. *Am J Lifestyle Med* 2018;12:148–59.
12 Holfelder B, Schott N. Relationship of fundamental movement skills and physical activity in children and adolescents: a systematic review. *Psychol Sport Exerc* 2014;15:382–91.
13 Lubans DR, Morgan PJ, Cliff DP, *et al.* Fundamental movement skills in children and adolescents: review of associated health benefits. *Sports Med* 2010;40:1019–35.
14 Logan SW, Kipling Webster E, Getchell N, *et al.* Relationship between fundamental motor skill competence and physical activity during childhood and adolescence: a systematic review. *Kinesiology Rev* 2015;4:416–26.
15 Utesch T, Bardid F, Busch D, *et al.* The relationship between motor competence and physical fitness from early childhood to early adulthood: a meta-analysis. *Sports Med* 2019;49:541–51.
16 Westendorp M, Houwen S, Hartman E, *et al.* Are gross motor skills and sports participation related in children with intellectual disabilities? *Res Dev Disabil* 2011;32:1147–53.
17 Eguia KF, Capio CM, Simons J. Object control skills influence the physical activity of children with intellectual disability in a developing country: the Philippines. *J Intellect Dev Disabil* 2015;40:265–74.
18 Einarsson IO, Olafsson A, Hinriksdottir G, *et al.* Differences in physical activity among youth with and without intellectual disability. *Med Sci Sports Exerc* 2015;47:411–8.
19 McCoy SM, Jakicic JM, Gibbs BB. Comparison of obesity, physical activity, and sedentary behaviors between adolescents with autism spectrum disorders and without. *J Autism Dev Disord* 2016;46:2317–26.
20 Wouters M, Evenhuis HM, Hilgenkamp TIM. Physical fitness of children and adolescents with moderate to severe intellectual disabilities. *Disabil Rehabil* 2019;1–11.
21 Maiano C, Hue O, Morin AJ, *et al.* Prevalence of overweight and obesity among children and adolescents with intellectual disabilities: a systematic review and meta-analysis. *Obes Rev* 2016;17:599–611.
22 O' Shea M, O' Shea C, Gibson L, *et al.* The prevalence of obesity in children and young people with down syndrome. *J Appl Res Intellect Disabil* 2018;31:1225–9.
23 Pratt HD, Greydanus DE. Intellectual disability (mental retardation) in children and adolescents. *Prim Care* 2007;34:375–86. abstract ix.
24 World Health Organization. Definition: intellectual disability. *World Health Organization*. 2014. Available http://www.euro.who.int/en/health-topics/noncommunicable-diseases/mental-health/news/news/2010/15/childrens-right-to-family-life/definition-intellectual-disability (accessed 15th Jul 2020)
25 Faras H, Al Ateeqi N, Tidmarsh L. Autism spectrum disorders. *Ann Saudi Med* 2010;30:295–300.
26 World Health Organization. Autism spectrum disorders. *World Health Organization*. 2019. Available https://www.who.int/news-room/fact-sheets/detail/autism-spectrum-disorders (accessed 15th Jul 2020)
27 Thurm A, Farmer C, Salzman E, *et al.* State of the field: differentiating intellectual disability from autism spectrum disorder. *Front Psychiatry* 2019;10:526.
28 Totsika V, Hastings RP, Emerson E, *et al.* A population-based investigation of behavioural and emotional problems and maternal mental health: associations with autism spectrum disorder and intellectual disability. *J Child Psychol Psychiatry* 2011;52:91–9.
29 Burton AW, Miller DE. *Movement skill assessment*. Champaign, IL: Human Kinetics, 1998.
30 Robertson S, Kremer P, Aisbett B, *et al.* Consensus on measurement properties and feasibility of performance tests for the exercise and sport sciences: a Delphi study. *Sports Med Open* 2017;3:2.
31 Mokkink LB, Terwee CB, Patrick DL, *et al.* The COSMIN study reached international consensus on taxonomy, terminology, and definitions of measurement properties for health-related patient-reported outcomes. *J Clin Epidemiol* 2010;63:737–45.
32 Beattie M, Murphy DJ, Atherton I, *et al.* Instruments to measure patient experience of healthcare quality in hospitals: a systematic review. *Syst Rev* 2015;4:97.
33 Klingberg B, Schranz N, Barnett LM, *et al.* The feasibility of fundamental movement skill assessments for pre-school aged children. *J Sports Sci* 2019;37:378–86.
34 Cools W, Martelaer KD, Samaey C, *et al.* Movement skill assessment of typically developing preschool children: a review of seven movement skill assessment tools. *J Sports Sci Med* 2009;8:154–68.
35 Griffiths A, Toovey R, Morgan PE, *et al.* Psychometric properties of gross motor assessment tools for children: a systematic review. *BMJ Open* 2018;8:e021734.
36 Scheuer C, Herrmann C, Bund A. Motor tests for primary school aged children: a systematic review. *J Sports Sci* 2019;37:1097–112.
37 Hulteen RM, Barnett LM, True L, *et al.* Validity and reliability evidence for motor competence assessments in children and adolescents: a systematic review. *J Sports Sci* 2020;1–82.
38 Bardid F, Vannozzi G, Logan SW, *et al.* A hitchhiker's guide to assessing young people's motor competence: deciding what method to use. *J Sci Med Sport* 2019;22:311–8.
39 Logan SW, Barnett LM, Goodway JD, *et al.* Comparison of performance on process- and product-oriented assessments of fundamental motor skills across childhood. *J Sports Sci* 2017;35:634–41.

40 Berkeley SL, Zittel LL, Pitney LV, *et al.* Locomotor and object control skills of children diagnosed with autism. *Adapted Phys Act Q* 2001;18:405–16.
41 Block ME, Taliaferro AM. Assessing children with autism in physical education. In: Grenier M, ed. *Physical education for children with autism*. Champaign, IL: Human Kinetics, 2014.
42 Richards SB, Brady MP, Taylor RL. *Cognitive and Intellectual disabilities: historical perspectives, current practices, and future directions*. 2nd edn. NY: Routledge, 2015.
43 Wilson RB, Enticott PG, Rinehart NJ. Motor development and delay: advances in assessment of motor skills in autism spectrum disorders. *Curr Opin Neurol* 2018;31:134–9.
44 Gkotzia E, Venetsanou F, Kambas A. Motor proficiency of children with autism spectrum disorders and intellectual disabilities: a review. *Eur Psychomotricity J* 2017;9:46–69.
45 Fournier KA, Hass CJ, Naik SK, *et al.* Motor coordination in autism spectrum disorders: a synthesis and meta-analysis. *J Autism Dev Disord* 2010;40:1227–40.
46 Staples KL, MacDonald M, Zimmer C. Assessment of motor behavior among children and adolescents with autism spectrum disorder. *Int Rev Res Dev Disabilities* 2012;42:179–214.
47 Maiano C, Hue O, April J. Fundamental movement skills in children and adolescents with intellectual disabilities: a systematic review. *J Appl Res Intellect Disabil* 2019;32:1018–33.
48 Capio CM, Eguia KF, Simons J. Test of gross motor development-2 for Filipino children with intellectual disability: validity and reliability. *J Sports Sci* 2016;34:10–17.
49 Samsudin NA, Ahmad AC. The fundamental movement skills (FMS) of children with autism spectrum disorder (ASD) and typical development (TD) children using the test of gross motor development (TGMD-2) in Penang, Malaysia. *Int J Manage Dev Stud* 2018;7:8–17.
50 Mache MA, Todd TA. Gross motor skills are related to postural stability and age in children with autism spectrum disorder. *Res Autism Spectr Disord* 2016;23:179–87.
51 Jeoung B. Motor proficiency differences among students with intellectual disabilities, autism, and developmental disability. *J Exerc Rehabil* 2018;14:275–81.
52 Prinsen CAC, Mokkink LB, Bouter LM, *et al.* COSMIN guideline for systematic reviews of patient-reported outcome measures. *Qual Life Res* 2018;27:1147–57.
53 Moher D, Liberati A, Tetzlaff J, *et al.* Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *PLoS Med* 2009;6:e1000097.
54 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.
55 Mokkink LB, de Vet HCW, Prinsen CAC, *et al.* COSMIN risk of bias checklist for systematic reviews of patient-reported outcome measures. *Qual Life Res* 2018;27:1171–9.
56 Ulrich DA. *Test of gross motor development*. 3rd edn. Austin, TX: PRO-ED, 2019.
57 Bruininks RH, Bruininks BD. *Bruininks-Oseretsky test of motor proficiency*. 2nd edn. BOT-2. Minneapolis, MN: Pearson, 2005.
58 Henderson SE, Sugden DA, Barnett AL. *Movement assessment battery for children-2 second edition (movement ABC-2)*. London, UK: The Psychological Corporation, 2007.
59 Folio MR, Fewell R. *Peabody developmental motor scales, second edition (PDMS-2)*. Austin, TX: Pearson, 2000.
60 Kalgotra R, Warwal JS. Development and standardization of a test of motor proficiency in children with intellectual disabilities in India. *Disability CBR Inclusive Dev* 2017;28:80–97.
61 Cratty BJ, Martin MM. *Perceptual motor efficiency in children*. Philadelphia, PA: Lea & Febiger, 1969: 183–94.
62 Dunn JM, Morehouse JW, Dalke B. *Game, exercise and leisure sport curriculum*. Corvallis: Oregon State University, Department of Physical Education, 1979.
63 Bricker D, Squires J. *Ages & stages questionnaire: a parent-completed, child monitoring system*. 2nd edn. Paul H Brookes Publishing Co, 1999.
64 Dite W, Temple VA. A clinical test of stepping and change of direction to identify multiple falling older adults. *Arch Phys Med Rehabil* 2002;83:1566–71.
65 Capio CM, Mak TCT, Tse MA, *et al.* Fundamental movement skills and balance of children with Down syndrome. *J Intellect Disabil Res* 2018;62:225–36.
66 Schott N, Holfelder B. Relationship between motor skill competency and executive function in children with Down's syndrome. *J Intellect Disabil Res* 2015;59:860–72.
67 Simons J, Daly D, Theodorou F, *et al.* Validity and reliability of the TGMD-2 in 7–10-year-old Flemish children with intellectual disability. *Adapt Phys Activ Q* 2008;25:71–82.
68 Allen KA, Bredero B, Van Damme T, *et al.* Test of gross motor development-3 (TGMD-3) with the use of visual supports for children with autism spectrum disorder: validity and reliability. *J Autism Dev Disord* 2017;47:813–33.
69 Pitetti K, Miller RA, Loovis M. Balance and coordination capacities of male children and adolescents with intellectual disability. *Adapt Phys Activ Q* 2017;34:1–18.
70 Wuang YP, Su CY. Reliability and responsiveness of the Bruininks-Oseretsky test of motor proficiency-second edition in children with intellectual disability. *Res Dev Disabil* 2009;30:847–55.
71 Wuang YP, Lin YH, Su CY. Rasch analysis of the Bruininks-Oseretsky test of motor proficiency-second edition in intellectual disabilities. *Res Dev Disabil* 2009;30:1132–44.
72 Wuang YP, Su CY, Huang MH. Psychometric comparisons of three measures for assessing motor functions in preschoolers with intellectual disabilities. *J Intellect Disabil Res* 2012;56:567–78.
73 Vanvuchelen M, Van Schuerbeeck L, Braeken MA. Screening accuracy of the parent-completed ages and stages questionnaires - second edition as a broadband screener for motor problems in preschoolers with autism spectrum disorders. *Autism* 2017;21:29–36.
74 Roswal PM, Sherrill C, Roswal GM. A comparison of data based and creative dance pedagogies in teaching mentally retarded youth. *Adapted Phys Act Q* 1988;5:212–22.
75 Verma A, Samuel AJ, Aranha VP. The four square step test in children with Down syndrome: reliability and concurrent validity. *J Pediatr Neurosci* 2014;9:221–6.
76 Williams HG, Pfeiffer KA, Dowda M, *et al.* A field-based testing protocol for assessing gross motor skills in preschool children: the champs motor skills protocol (CMSP). *Meas Phys Educ Exerc Sci* 2009;13:151–65.
77 Tyler R, Foweather L, Mackintosh KA, *et al.* A dynamic assessment of children's physical competence: the dragon challenge. *Med Sci Sports Exerc* 2018;50:2474–87.
78 Herrman C, Seelig H. MOBAK - 5–6 basic motor competencies in fifth and sixth grade. 2018. Available http://mobak.info/wp-content/uploads/2018/06/MOBAK_5-6_engl.pdf (accessed 15th Jul 2020)
79 Bisi MC, Pacini Panebianco G, Polman R, *et al.* Objective assessment of movement competence in children using wearable sensors: an instrumented version of the TGMD-2 locomotor subtest. *Gait Posture* 2017;56:42–8.
80 Bruininks RH. *Bruininks-Oseretsky test of motor proficiency - owner's manual*. Circle Pines, MN: American Guidance Service, 1978.