

Movement skill assessment of typically developing preschool children: A review of seven movement skill assessment tools

Wouter Cools¹✉, Kristine De Martelaer¹, Christiane Samaey¹ and Caroline Andries²

¹ Department of Movement Education and Sport Training, Faculty of Physical Education and Physiotherapy, ² Department of Developmental and Lifespan Psychology, Faculty of Psychology and Education, Vrije Universiteit Brussel, Brussel, Belgium

Abstract

The importance of movement is often overlooked because it is such a natural part of human life. It is, however, crucial for a child's physical, cognitive and social development. In addition, experiences support learning and development of fundamental movement skills. The foundations of those skills are laid in early childhood and essential to encourage a physically active lifestyle. Fundamental movement skill performance can be examined with several assessment tools. The choice of a test will depend on the context in which the assessment is planned. This article compares seven assessment tools which are often referred to in European or international context. It discusses the tools' usefulness for the assessment of movement skill development in general population samples. After a brief description of each assessment tool the article focuses on contents, reliability, validity and normative data. A conclusion outline of strengths and weaknesses of all reviewed assessment tools focusing on their use in educational research settings is provided and stresses the importance of regular data collection of fundamental movement skill development among preschool children.

Key words: Early childhood, psychomotor performance, motor development, validity, reliability.

Introduction

Usually, children attending preschool range in age from three to six, although in Europe some differences between countries exist (Eurydice, 2002). This age period is a sensitive period for the development of fundamental movement skills [FMS] (Gallahue and Donnelly, 2003). Because most preschool children are naturally curious, love to play and explore, these FMS are learned very easily. Especially when stimulation, opportunities to play and to be physically active or sport are offered. The mastery of certain FMS is a prerequisite for daily life functioning and participation in later physical or sport-specific activities.

At an early age, gross movement skills are necessary to move, stabilize and control body and objects while exploring the environment. Later in life, well developed gross movement skills help individuals to function more smoothly. Fine movement skills are necessary for the development of basic self-help skills. Also drawing and writing are based on fine movement skill development. Later in life well developed fine movement skills are as important as gross movement skills.

During infancy, development is evaluated almost

exclusively by motor development (Berk, 2003). Once a child can reach, grasp and walk, however, interest in the further development of more complex movement skills is reduced and more attention is given to the development of cognitive, social and emotional aspects. Motor development is basically only taken into consideration when dysfunctions or inefficient movement behavior appears (Davies, 2003). Research in the area of movement skill development mainly focuses on motor impairment and motor deficits. Hence, research on FMS development and performance in developing children is scarce and rather fragmentary. The information which is available is mostly based on the sequences of developmental change in movement patterns and can be found in literature such as Gallahue and Ozmun (2006) and Haywood and Getchell (2005). Normative data on FMS development and performance are mainly derived from control or normative samples in research studies on children with a developmental disorder. In general, normative data on FMS development and performance of European samples of preschool are scarce.

Different tools to assess movement performance in early childhood are available (Barnett and Peters, 2004; Simons, 2004; Vallaey and Vandroemme, 2001; Wiart and Darrah, 1999). Most of these tools are aimed at a specific target group and hence have specific content. The movement assessment can be norm- or criterion referenced. A norm-referenced test compares the child's performance to that of a normative group, and quantifies the child's movement skill competence. A criterion-referenced test compares the child's performance to predetermined criteria. A criterion-referenced test takes into account the qualitative aspects of the movements required to perform the movement skill item. A second form of movement skill assessment is through pupil monitoring instruments and is mainly used by teachers. Although many child monitoring instruments (SIG, 2005; Bertrands et al., 2003; van Gelder and Stroes, 2002) exist, there is little agreement on what might be expected in relation to children's FMS development (Haywood and Getchell, 2005).

In addition to earlier reviews such as Barnett and Peters (2004), Tieman et al. (2005), Yoon et al. (2006) and Wiart and Darrah (2001), this article provides a review of seven movement skill assessment tools with the scope on movement development and performance in typical preschool children.

The review explores the potential usefulness to

assess movement performance in an educational research context. Six of these tests are often used or referred to in a European and international context (Bös, 2003; Simons, 2004; Vallaey and Vandroemme 1999). The seventh, the Maastrichtse Motoriek Test (MMT) is a recently developed tool and is added because of its innovative combination of quantification and qualification of movement skills development and performance. Different aspects of the assessment tools are described, including content and test administration, validity, reliability, and normative data. The review includes the following tests:

- Motoriktest für Vier- bis Sechsjährige Kinder (MOT 4-6),
- Movement Assessment Battery for Children (Movement-ABC),
- Peabody Development Scales (PDMS),
- Körperkoordinationstest für Kinder (KTK),
- Test of Gross Motor Development (TGMD),
- the Maastrichtse Motoriek Test (MMT),
- the Bruininks-Oseretsky test of Motor Proficiency (BOTMP).

Although the movement skill assessment tools vary in specific applications, the basic concepts of assessment all operate similarly.

Movement skill assessment tools

Motoriktest für vier- bis sechsjährige Kinder (MOT 4-6) [Zimmer and Volkamer, 1987]

The MOT 4-6 test is of German origin and has been developed to contribute to the assessment of FMS development. In addition, the tool creates an opportunity for early detection of FMS delay or deficiency. The test is rooted in both the Lincoln Oseretsky Motor Development Scales (LOMDS) and the Körperkoordinationstest für Kinder (KTK) to which adaptations have been made to make the test appropriate for the specific age group of preschool children (Zimmer and Volkamer, 1987). The authors believe that children in this age group have specific needs and require a different pedagogical approach. Therefore, the age range (4 to 6 year-olds) applicable for this tool is kept very narrow. The test features 18 different items including locomotion, stability, object control and fine movement skills (see Table 1). The standardized manual comprises exact descriptions of every item: detailed task description, required material, indications on important aspects, specific simple instructions for the child and a three-point rating scale from 0 (skill not mastered) to 2 (skill mastered). A well-organized score sheet enhances standardization. Additionally, free space for qualitative notes about the child or its performance is provided. To guarantee maximum attractiveness of the test, the subsequent items have different motor demands. The total test time per child fluctuates between 15 and 20 minutes. As some assignments have to be performed barefoot, test time might be slightly prolonged. For some children this might even be experienced as a barrier.

In accordance with the test purpose, the total motor score expresses children's FMS performance. The MOT 4-6 is product-oriented and refers to a norm. Half-yearly norms are derived from a sample of 548 typically devel-

oping German preschool children. No separate normative data for boys and girls were included because of an absence of significant gender differences in total motor scores. Qualified test administrators have to be familiar with every test item's specific instructions and should be able to demonstrate every task adequately.

The MOT 4-6 is a coordination assessment tool for preschool children (Bös, 2003), recommended for educational research purposes because of its specific age range (Vallaey and Vandroemme, 1999). A test revision of the MOT 4-6 is in progress and the age range will be extended to 8 years (Zimmer, 2006).

Movement Assessment Battery for Children (Movement-ABC – Movement-ABC 2) (Henderson and Sugden, 1992; Henderson, Sugden and Barnett 2007)

The Movement-ABC assesses the developmental status of FMS; with a focus on detection of delay or deficiency in a child's movement skill development (Vallaey and Vandroemme, 1999). The Movement-ABC test is a revision of the Test of Motor Impairment (TOMI) and originates from the Oseretsky scales for the motor capacity of children (Simons, 2004; Burton and Miller, 1998). The test is suitable for children between 4 and 12 years of age and consists of 32 items, subdivided into 4 age bands. Each age band includes 8 individual test items measuring movement skills in three categories: manual dexterity skills, ball skills and balance skills. Taking the test requires 20 to 30 minutes. A total impairment score expresses the child's test performance. Each item is rated on a 6-point rating scale, where 5 equates to the weakest performance and 0 equals the best performance. Profile scores provide more specific information on the child's movement skill performance of each individual category. Qualitative observations are optional (Henderson and Sugden, 1992).

The most important advantages of the test are: its availability in several European countries, its cross cultural validity which is based on comparison with local sample data (e.g. Smits-Engelsman, 1998; Petermann, 2008; Soppelsa and Albaret, 2004) and its simple test administration, which facilitates large sample screening over a short period. The disadvantages of the test are: its rather large age range (loss of specificity) and its unfavorable proportion of test items versus time required for test administration (8 items/20-30 min). Unlike other movement skill tests, such as BOTMP, which measures the child's strengths and weaknesses over a wide range of skills, the Movement-ABC is limited to the movement skills of a certain age band. The Movement-ABC is a product-oriented test and refers to a norm. In the revised version (Henderson et al., 2007) qualitative observations have been added. However, they do not have an impact on the score and are meant to specify the difficulties that children encounter when performing a movement skill task. Following the Movement ABC checklist, the age range of Movement-ABC 2 checklist is extended (ages 5 to 12) and focuses on how a child manages everyday tasks encountered in school and at home. The checklist has a motor and a non-motor component that provides information on direct and indirect factors that might affect movement. The checklists, however, are outside the scope

of this article and will not be discussed in detail.

The test is used as a screening instrument for problems in the development of integrated motor skills (Rosenbaum et al., 2004; Van Waelvelde et al., 2004). According to these authors, the tool is especially useful in exploring issues in the functional integration of motor control or problems that often appear for the first time in late preschool and early primary school years. Burton and Miller (1998) consider the test suitable for assessment of motor abilities, early milestones, FMS and specialized movement skills. Cross cultural validity (with/without modifications of the test) has been supported by a number of studies and resulted in the Movement-ABC test being translated in several languages (e.g. Chinese, Dutch, Danish, Swedish, Italian and Japanese [Barnet and Peters, 2004; Chow et al., 2001; Chow et al., 2006]).

The revised version (Henderson et al., 2007) includes the following main points:

- The test is divided over three components: a standard test, a checklist and a companion manual which describes an ecological approach on intervention for children with movement difficulties.
- Age extension (from 3 to 16 years) and reorganization of age bands (3 to 6, 7 to 10 and 11 to 16 years).
- Revision of the test content:
 - material (manufactured in plastic instead of wood to overcome inaccurate measuring as a result of wear and tear),
 - tasks: individual item changed; e.g. bicycle trial has changed into a drawing trial; rolling ball into goal has changed into throwing a beanbag onto a mat; etc. All of these changes have been made with the intention of increasing correspondence, sensitivity and consistency between test items of the different age bands;
 - instruction: test instructions were clarified to reduce ambiguity in test administration and scoring.
- Updated normative data: 1172 children participated in the study between November 2005 and July 2006 (age band 1 (3-6 y) n = 431; age band 2 (7-10 y) n = 333 and age band 3 (11-16 y) n = 408). All data were gathered in Britain and Northern Ireland and the sample was stratified for geographic region, population density, social class, and race or ethnicity.
- Additional normative data collection on children from outside the UK is in progress.

Peabody Developmental Motor Scales- Second Edition (PDMS-2) [Folio and Fewell, 1983; Folio and Fewell, 2000]

The PDMS-2 is a movement skill assessment tool that measures gross and fine movement skills. It focuses on assessment and intervention or treatment programming for children with disabilities. The test manual states that the test estimates a child's motor competence relative to his or her peers, determines the balanced development of fine and gross motor movement skills, identifies skill deficits and evaluates progress. Therefore, it can be used as a research tool.

The PDMS-2 is a revision of the original PDMS published in 1983. It consists of 6 subtests of which 4 involve gross and 2 involve fine movement skills. The test is designed to assess movement skills of children

from birth to 6 years of age. The gross movement subtests include: reflexes (8 items), stationary performances (30 items), locomotion (89 items) and object manipulation (24 items). The fine movement subtests include: grasping (26 items) and visual-motor integration (72 items).

According to the authors (Folio and Fewell, 2000), the PDMS-2 is a standardized instrument including reliable and valid scales. The test is standardized on a sample of children from 20 states across the United States. Maturation of the body as well as environmental factors (e.g. race and ethnicity of the children) is considered in the sampling.

The total motor score is the sum of all 6 subtest scores. The test uses a 3-point rating system of which 2 equals an attained skill, 1 a developing skill and 0 a non acquired skill. This offers the opportunity to measure progress later. Every item includes criteria for each rating. The PDMS-2 is primarily used for individual assessment, but adaptations have been made to allow group assessments (Simons, 2004). Administering the whole test varies between 45 and 60 minutes.

The PDMS-2 test shows several improvements (Simons, 2004; Vanvuchelen et al., 2003). Firstly, normative data have been expanded to 2003 American and Canadian children (Data collected between 1997-1998). The authors of the PDMS-2 report that reliability and validity have been thoroughly analyzed and optimized (Folio and Fewell, 2000). Also new score criteria have been added to the initial PDMS. Illustrations have been included to clarify assignments. Some items have been adapted and activity cards changed into motor activity programs. Correlations of $r = 0.84$ are established between PDMS and PDMS-2. A clear structure is obtained through factor analysis and is thoroughly described in the test manual.

The test discriminates motor developmentally delayed and disordered children from typically developing children. Interchangeable and cross cultural use of the PDMS and PDMS-2 are not recommended without precaution and proper adaptation. Darrach et al. (2007) found no evidence for equivalent use of PDMS and PDMS-2 in 4 year-olds. According to Tripathi et al (2008), it is not possible to develop culturally sensitive assessment tools across geographical regions and environments. Therefore, it is necessary to evaluate cultural sensitivity of the assessment tool when using it in a particular region (sample of children from India). Vanvuchelen et al. (2003) support this statement by concluding that for Flemish children the use of the PDMS-2 (with American normative data) was not reliable enough to distinguish between children suffering from motor developmental delay or disorders and typically developing children. The PDMS-2 overestimates the 5-year-old Flemish preschool child. According to Vanvuchelen et al. (2003), the PDMS-2 is consistent enough to conclude on a child's general FMS developmental status. However, more thorough standardization of the test is needed because of the dependence on observer interpretation. The PDMS-2 is process, as well as, product-oriented and refers to a criterion as well as to a norm (Vallaey and Vandroemme, 1999).

Because the content has been virtually unchanged, Burton and Miller's (1998) conclusion on PDMS suitability for the assessment of motor abilities, fundamental

motor skills and early milestones still holds. The Peabody Developmental Scales at 3 and 4 years of age can screen particularly for the appearance/onset of problems in the development of integrated motor skills, in the face of what appear to be adequate gross motor abilities (Rosenbaum et al., 2004). Vanvuchelen et al. (2003) recommend the use of PDMS-2 to determine the strengths and weaknesses of individual children in therapy planning.

Körperkoordinationstest für Kinder (KTK) [Kiphard and Shilling, 1974; Kiphard and Schilling, 2007]

The KTK is appropriate for children with a typical developmental pattern, as well as for children with brain damage, behavioral problems or learning difficulties. The test assesses gross body control and coordination, mainly dynamic balance skills. The KTK is a shortened version (from 6 to 4 items) of the Hamm-Manburger Körperkoordination Test für Kinder of Kiphard and Schilling (1974). The test covers an age range from 5 to 14. Assessing one child takes approximately 20 minutes.

The test is thoroughly standardized and considered highly reliable (Valaey and Vandroemme, 1999). It is easy to set up and takes little time to administer. This results in rapid screening of the balance function. The test items, however, are not learned quickly, so the test can be used for evaluating therapy and interventions. The KTK has separate normative data tables for boys and girls for 2 it's items. The test is limited to one aspect of gross movement skill assessment, object control and locomotion functioning are not integrated in the test (see Table 2). The KTK-test is a product-oriented test that refers to a norm. In spite of the KTK lasting for a long time, its value is preserved. Especially when one is specifically interested in the evaluation of balance skill development among children, the KTK offers a highly reliable and standardized opportunity for assessment (Gheysen et al., 2008). Furthermore the test is still used for the criterion validity studies of other assessment tools, e.g. M-ABC 2 (Henderson et al., 2007).

Test of Gross Motor Development, Second Edition (TGDM-2) [Ulrich,1985; Ulrich, 2000]

The TGDM-2 measures gross movement performance based on qualitative aspects of movement skills. According to the author, the test can be used to identify children who are significantly behind their peers in gross motor performance, to plan programs to improve skills in those children showing delays and to assess changes as a function of increasing age, experience, instruction or intervention. The TGDM-2 is a revision of the original Test of Gross Motor Development (TGMD), published in 1985 (Ulrich, 1985). The age range (3 to 10 years) covers the period in which the most dramatic changes in a child's gross movement skill development occur (Ulrich, 2000). The test includes locomotion and object control skills. The locomotion part consists of six consecutive items: running, galloping, hopping, leaping, horizontal jumping and sliding. The object control subtest consists of six consecutive items: two-hand striking a stationary ball, stationary dribbling, catching, kicking, overhand throwing and underhand rolling. The child has to perform every

item twice. When the performance is correct a score of 1 is marked, incorrect performances are scored 0. The sum of both performances represents the final score for each item. Standard scores for both locomotion and object control parts can be calculated and age equivalents can be derived. The test is administered in 15 to 20 minutes and requires equipment that is commonly used during PE.

The test revision shows several improvements. Ulrich (2000) reports on reliability and validity issues which have been thoroughly revised: internal consistency and stability coefficients have been added and reliability coefficients have been computed for subgroups of the normative sample, validity for a wide variety of subgroups has been obtained. New normative data from the USA were gathered. The normative sample has undergone specific changes: the sample was stratified (by age, relative to geography, race, gender and residence); norms fall into half-year periods and gender normative tables were created for the subtest object control. Some test items changed, pictures were redrawn and the skipping item was dropped and underhand rolling was added to the subtest object control.

A great advantage of the TGDM-2, in addition to performance assessment, is the incorporation of qualitative aspects in the assessment. Regrettably, no stability subtest is included. Simons and Van Hombecck (2003) conclude that Flemish children score significantly lower than American children on the TGDM-2. The cultural differences as a possible explanation for this underachievement are proposed. The object control items, especially the striking and overhand throwing items (both highly related to baseball skills), might be inappropriate to use cross culturally as a standard for object control assessment. The TGDM-2 is a process and product-oriented test that refers to a criterion and a norm. Because no extensive content changes have been made, Burton and Miller's conclusion (1998) on the TGMD's suitability to assess motor abilities and FMS still holds for TGDM-2.

Maastrichtse Motoriek Test (MMT) [Vles et al., 2004]

Vles et al. (2004) recently designed a new assessment tool, the Maastrichtse Motoriek Test (MMT). The purpose of the MMT is to objectively assess qualitative aspects of movement skill patterns in addition to quantitative movement skill performance. The test distinguishes between children with and without normal motor behavior. The authors claim to detect children at risk for Attention Deficit Hyperactivity Disorder (ADHD) at an early age. The MMT measures fine as well as gross movement skills. The test is suitable for 5 to 6-year-old children, the age period seen as the transition stage between pre- and primary school. The MMT includes 70 items of which 34 measure quantitative and 36 measure qualitative aspects of movement skill performance. To score the child's performance on an item, a three-point scale is used; from 0 to 2. It takes 20-25 minutes to administer the test. Scoring qualitative aspects of movement requires well trained observation skills. Therefore, next to a clear description, images and small video fragments (showing a weak, a moderate and a good performance) for all qualitative test items have been included on a CD-Rom. A group of 487 children in the second year of elementary school in the

Netherlands participated in the normative data sample. To contribute to the content validity a panel of experts cooperated closely in the development of the MMT. The MMT observations of two well-trained observers were compared with a school doctor's (with > 20 years of experience) judgment on the children's motor development (normal versus abnormal). The school doctor assessed the children separately and was unaware of the MMT test results. Video recordings of 24 children were made during testing to investigate intra-rater reliability. The test-retest reliability intra-class correlation coefficients (ICC) ranged from $r = 0.43$ to 0.93 . A group of children ($n=43$) was tested twice by the same examiner, the children scored slightly better on the second trial but no significant differences were found. Two raters scored 42 children independently at the same time, ICCs of inter-rater reliability varied from $r = 0.92$ to 0.97 . Areas under curve were calculated, and varied from $r = 0.81$ to 0.86 . Intra-rater reliability ICCs varied from $r = 0.72$ to 0.98 . Separate normative data tables for boys and girls are provided because of significant gender differences.

A promising strength of the MMT is that it includes qualitative observations in the total movement skill score and thus provides a more holistic view on the child's strengths and weaknesses. Among the limitations are the absence of locomotor skill items and the very small age range that is covered.

Bruininks-Oseretsky Test of Motor Proficiency (BOTMP-BOT-2) [Bruininks, 1978; Bruininks and Bruininks, 2005]

The Bruininks-Oseretsky Test of Motor Proficiency (BOTMP) and its review the Bruininks-Oseretsky Test of Motor Proficiency, second edition (BOT-2) are tools to assess fine and gross movement skill development. They are used to identify individuals with mild to moderate motor coordination deficits. The test is suitable for individuals aged 4 to 21 years. The complete BOT-2 features 53 items and is divided into 8 subtests: fine motor precision (7 items), fine motor integration (8 items), manual dexterity (5 items), bilateral coordination (7 items), balance (9 items), running speed and agility (5 items), upper limb coordination (7 items), strength (5 items). The items in every subtest become progressively more difficult. A short form of the BOT-2 can be used as a screening tool to achieve rapid and easy scoring reflecting overall motor proficiency. The BOT-2 Short Form comprises a subset of 14 items of the BOT-2 Complete Form and was constructed from data gathered in standardization (Bruininks and Bruininks, 2005). The Short Form features items from all subtests. A high correlation ($\sim r = 0.80$ s) was found between the short and long form of the BOT-2.

The selection of the items was based on the following criteria:

- to provide a broad and general view on the movement skill development status of a child;
- to represent significant aspects of motor behavior;
- to emphasize motor activity;
- to provide the opportunity to discriminate between a broad range of motor abilities;
- to fall within the possibilities of mild and moderately retarded children;

- to appeal to limited memory capacity and vocabulary of the child;
- material has to be easily transported.

The scoring system varies according to the individual items; it ranges from a 2-point scale to a 13-point scale. The raw scores can be converted into a standard numerical score. Results can be aggregated into a fine manual control composite, a manual coordination composite, a body coordination composite and a strength and agility composite. The sum of scores results in a total motor composite. The time required to assess one individual varies between 45 to 60 minutes for the complete test and between 15 and 20 minutes for the short form.

The revision goals included quality improvement of kit equipment, improvement of item presentation, improvement of measurements on the youngest children (4- and 5-year olds), improvement of functional relevance, expansion coverage of fine and gross motor skill and extension of age norms to the age of 21. The use of the test is recommended for motor impairment diagnosis, screening, placement decisions, development and evaluation of motor training programs and supporting research goals. The BOTMP is frequently used in adapted PE, occupational therapy and physical therapy (Burton and Miller, 1998). According to Rosenbaum et al. (2004), the BOTMP is designed for assessment of motor skills in children, for those with motor dysfunctions in particular. Bruininks and Bruininks (2005) proved test validity for BOT-2 for individuals with developmental coordination disorder (DCD), mild to moderate mental retardation (MR), and high-functioning autism /Asperger's Disorder.

A total of 1520 children from 239 settings of all states in the US were included in the collection of normative data in the beginning of 2005.

Peerlings (2007) lists some important barriers for the use of this test:

- the test is only obtainable by medical and paramedical professions, and even then it is not very easy to obtain the assessment battery.
- Because of the adaptation of some items a more thorough training is required and 18 m of running space is required for the test setting.
- The order of test items on the scoring sheet do not comply with the order of subtests assessment.
- For some of the younger children the time required to complete the test is too long, so it is recommended to spread the assessment over two test sessions.

Table 1 provides a summary and overview of all assessment tools that have been described previously. An overview of the content of the movement skill items of each test are shown in Table 2.

Administrative and organizational aspects

Test Choice

Table 1 shows administrative aspects of the reviewed tools and can be used to guide the selection of the most appropriate one for a specific educational research goal or clinical purpose. Adequate standardized tools include following basic criteria: clear conversion tables in the test manual, an all-inclusive test kit, clear descriptions of test

Table 1. Administrative aspects of the motor assessment tools.

	Purpose	Assessment / training	Age (y:m)	Assessment time (min)	Number of items	Required equipment	Raw score conversion	Cost 2006*
MOT 4-6	Assesses motor developmental status at pre-school age	General assessment of gross and fine motor skills	4:0-6:11	15-20	18 items	Test kit, stopwatch and clipboard	Percentile rank, T-score, C-score, DMQ Stanine	418€
Movement ABC	Identify and describe motor impairments in daily life	SMD, Level measurement Evaluation of treatment (description)	4:0 - 12:0	20-30	32 items: 4 age bands (4x8 items)	Test kit Clipboard and stopwatch	Percentile rank Total impairment score	954€
PDMS 2	MDA and programming for young children with disabilities	In-depth assessment and training or remediation of gross and fine motor skills	0:0-6:11	LV: 45-60; STV:20 – 30	249 items	Test kit and additional required materials	GMQ, FMQ, TMQ, Percentile, Standard scores, z-score, T-score, AE, scaled score, mean MAE	855€
KTK	Assesses general dynamic balance skill	Screening for children suffering from brain damage, behavioral and learning disturbances Screening gross motor skill (balance)	5:0 - 14:0	20	4 items	Balance beams (6, 4.5 and 3 cm); sponge blocs (5 - 60 cm); wooden slat, 2 wooden boxes	Percentile rank, Motor quotient	524€
TGMD-2	Identify children who are significantly behind their peers	Identify, plan, assess changes in relation to age or experience, assess changes after intervention or instruction	3:0 - 10:0	15 – 20	12 items	Masking tape, chalk, traffic cones, 10-15 cm light ball, 20-25 cm playground ball, 15-20 cm sponge ball, tennis ball, 20-25 cm plastic or slightly deflated playground ball, tape measure	Percentile rank, Standard scores, age equivalent, Gross Motor Quotient	262€
MMT	Objectify qualitative and quantitative aspects of movement	Detection of Attention Deficit Hyperactivity Disorder	5:0 – 6:11	LV: 30 SV: 7	LV: 70 items SV: 20 items	Manual, instruction CD-rom, score sheets, stopwatch, plastic ball, scotch tape	Percentile rank, Quality of movement	175€ + 272€ CD-rom Stand alone version
BOT-2	Identification of deficits in individual with light to moderate motor coordination problems	Profile analysis to evaluate an individual's strengths and weaknesses, Clinical validity studies on high-functioning autism/ Asperger's Disorder, developmental coordination disorder, and mild-to moderate mental retardation	4:0 - 21:0	LF: 45 - 60 SF: 15 - 20	LF: 53 items SF: 14 items	Test kit, and additional required equipment a tape measure, a stopwatch, 2 chairs, a table and a clip board	Subtest and Composite scores, Total Motor Composite; Standard Score, Scale score, Percentile Ranks, Standard deviations	1352€

* VAT including, SMD = Screening of motor difficulties, MDA = Motor development assessment, DMQ = Developmental Motor Quotient, GMQ = Gross Motor Quotient, FMQ = Fine Motor Quotient, TMQ = Total Motor Quotient, AE = Age Equivalent, MAE = Motor Age Equivalent

items and scoring instructions, a well organized score sheet and additional free space for additional qualitative information. Also assessment material should be easy to install and test items should be simple to instruct, demonstrate and easy to administer.

Assessment has to be performed in compliance with the standards for test users (Task Force on Standards

for Measurement in Physical Therapy TFSMPT, 1991), e.g. reporting test choice (e.g. describing practical use, physical setting, population and justification of test choice, comprising the reasons for not choosing a tool), test selection has to be based on what is best for the person being tested and ethical guidelines should be followed (e.g. oral consent from the assessed pre-school child,

Table 2. Movement skill content of assessment tools *.

Test	Gross Motor Movement Skills			Fine Motor Movement skills	Overlap
	Locomotion movement skill	Object control movement skill	Stability movement skill		
MOT 4-6	6	4	9	3	3
M-ABC	1	2	3	3	1
PDMS 2	16	10	18	18	2
KTK	0	/	4	/	0
TGMD-2	6	6	/	/	/
MMT	/	8	34	28	*
BOT-2	9	8	16	20	2

* This table is a concise summary of assessment tool item content. A detailed overview with specific information on individual items can be obtained from the authors.

informed consent from the child's parents, safety measures, etc.).

Test use

Special attention has to be paid to the test circumstances, because most tests have to take place 'on the spot'. Test manuals prepare for this environment by including clear descriptions of the physical requirements. Summarizing these guidelines in the test manuals an ideal test room should measure 6 m by 4 m, include one wall without openings or obstacles and a hard surfaced floor. The test room should be bright, quiet and well ventilated. Distraction of the tested child and loss of time is minimized when test items are set out around the room in advance.

Ideal test material should be standardized and appliances such as chairs and tables should be well adapted to the child's dimensions. Before starting the assessment, precautions should be taken to prevent dangerous situations.

Test raters

Another aspect of importance is the selection and training of examiners or test users. The test user's background should comply with the Standards for Test and Measurements in Physical therapy (TFSMPT, 1991). A test user's appropriate background includes basic knowledge of the theory and principles of tests and measurements, as well as standardized training and basic understanding of the child participant's background. Qualified test examiners should have read and studied the test materials and

Table 3. Reliability aspects.

Test	Inter-rater reliability	Intra-rater reliability	Test-retest reliability	Other reliability aspects
MOT 4-6	r = 0.88 (product-moment correlation) n = 32 children 5 raters	NR	r = 0.85 (4 weeks n = 47)	Split half r = 0.80 (Odd-even). Internal consistency: 0.806 (Cronbach's Alpha) SEM 2,185
M-ABC	Reliability reported for TOMI-H n = 360, 3 raters: ICC = 0.70 62-100% match between raters	NR	Reliability reported for TOMI-H n = 360, 3 raters: ICC = 0.75; 0.64 (4-6), 0.43 (6-8), 0.96 (9-10), 0.97 (11-12)	Reliability of cut-off scores, percentage of agreement by item, percentage of agreement for total impairment scores
PDMS 2	n = 60, r = 0.96, 2 raters	NR	2 groups of children: r = 0.89 (2-11 months; n = 20); r = 0.96 (12-17 months n = 30)	Standard error of measurement: ranging from 1-5. Internal consistency: 0.97 (Cronbach's Alpha). Coefficient alphas for six subscales over six age ranges from 0.71 to 0.98
KTK	> 0.85	ICC = 0.97 total (N = 68) ICC = 0.80 backward balance ICC = 0.95 sideward jump ICC = 0.94 displacing boxes ICC = 0.96 one leg jumping	r > 0.85. Comparison of averages between test and retest (all non significant)	Split half Reliability r = 0.92 - 0.97
TGMD-2	Correlation for subtests and Composites ranging from 0.97 - 0.99, r = 0.84 - 0.96, n = 60, 2 examiners scoring completed protocols	NR	Locomotion r = 0.85 Object Control: r = 0.88 Gross Motor Composite: r = 0.91	Internal consistency reliability coefficients averages between 0.85 - 0.91 (Cronbach's Alpha). Standard error of measurement (SEMs), Locomotion subtest: 1, Object Control subtest: 1, Gross Motor Quotient: 4-5. Reliability coefficients for subgroups: > 0.90s
MMT	n = 43, 0.92 Quantitative score, 0.97 Qualitative score, Individual scores all > 0.80 Range 0.83 - 0.97	Video tapes of 24 (n) children were taped during testing: video. Quantitative ICC = 0.72 to 0.98 (n = 24) Qualitative ICC = 0.82 - 0.95	n = 43 Ranging from 0.43 to 0.93	
BOT-2	Ranging from 0.92 to 0.99, n = 47 (Pearson's correlation), 2 raters	NR	n = 134, Mean subtest correlation coefficients ranging from 0.69 to upper 0.70s. Mean composite correlation coefficients ranging from 0.77 to low 0.80s. Short form and Total Motor Composite mid- to upper 0.80s	Subtest reliability ranging from high 0.70s to low 0.80s. Composite reliability coefficients ranging from high 0.80s to low 0.90s. Total Motor Composite mid 0.90s. SEMs of subtest scales near 2 SEMs of composite standard scores mid 3s

Table 4. Validity aspects.

Test	Construct	Content	Criterion related	Other
MOT 4-6	Test theoretical based test construction. Discriminates between typically developing and children with disabilities. No differences of total score for boys and girls. Enhancement correlates with age	Item analysis (n=2000)	Concurrent validity with KTK for 5-6 y olds: 0.78 (n = 181) after correction for age $r = 0.68$	Refers to a norm
Movement-ABC		NR	Concurrent validity with BOTMP $r = -0.53$ and with KTK $r = 0.62$	Refers to a norm Checklist refers to a criterion
PDMS 2	Scores correlate with age (ranging from 0.80 - 0.93). Group differentiation: Comparison of standard mean scores shows differences as one would expect. Confirmatory factor analysis ranging from low correlation 0.29 (grasping- sub items) to high correlation 0.89 (Gross Motor - Locomotion). Item validity supported by discriminating powers of TGDM-2	PDMS-2 1) Description of the rationale that underlies test item selection 2) Conventional item analysis of items by Item Response Theory 3) Differential functioning analysis 4) Item discrimination coefficients reported	Criterion prediction PDMS2 - PDMS: GMQ: $r = 0.84$, FMQ: $r = 0.91$ PDMS-2 and MSEL: > 0.80 for GMQ and FMQ, 0.73 - 0.83 for TMQ With the Mullen $r = 0.86$ (GM), $r = 0.80$ (FM)	Refers to a criterion and a norm
KTK	Construct validity proven by examining following assumptions: - absence of socio-cultural difference (urban and rural children) - gender differentiation in item normative data incorporated - differentiation for disadvantaged children (brain disorder, behavioral problems, mute) enhancement correlates with age	Factor analyses showed that the test evaluated dynamic body coordination and body control (dynamic balance)		Refers to a norm
TGDM-2	Ulrich addressed construct validity of TGDM-2 by examining five key assumptions - correlation with chronological age - differentiation between groups of individuals of average, below and above - items correlate highly with Total score of subtests - subtest composites correlate with each other - factor analysis: goodness of fit indexes ranging from 0.90 to 0.96	Three content experts judged whether skills selected are frequently taught in preschool and early elementary	Partial correlations: TGDM-2 and CSSA subtests: subtest Basic motor Generalization: 0.63 for Locomotion and 0.41 for Object Control Correlation between Composite and the CSSA subtest was 0.63	Refers to a criterion and a norm
MMT		Expert validity by a school doctor's judgment	AUC: range 0.80 - 0.87, ROC curves, cut-off points for sensitivity and specificity	Refers to a criterion and a norm
BOT-2	$r = 0.78$ (0.56 - 0.86) (test scores / chronological age) Discriminative validity with scores varying by whether children were developing typically or had varying grades of developmental difficulty	Theoretical and empirical sources of evidence include: test content, item fit and clinical groups.	BOTMP - BOT-2: Moderate to strong correlations for subtest and composite scores (0.45 - 0.73) ; high correlation for TMC and BC 0.80. PDMS-2 - BOT-2: TMC and TMQ strongly correlate (0.73); subtest and composite correlations are moderate to high (0.35 - 0.75). TVMS-R - BOT-2: correlation of Fine Motor Integration subtest is strong 0.74	Refers to a norm Goal directed activities

manual in advance. They then should practice administering and scoring until the specific rules and procedures are followed consistently and comfortably (Bruininks and Bruininks, 2005). When the results are reported, they should describe any incident that may have occurred during testing.

Reliability and validity

An assessment instrument that is not valid is utterly useless. An assessment tool that is not reliable cannot be valid. (Burton and Miller, 1998, p. 109)

This section includes an overview for each test individually. Reliability and validity as described in the test manuals are briefly summarized in Tables 3 and 4. To define magnitude of correlations Cohen's scale (1988), suggesting that a correlation of 0.5 is large, 0.3 is moder-

ate and 0.1 is small, was used. In general, reported internal consistency and inter-rater reliability coefficients are high for all tests. A general remark on concurrent validity is that test comparisons show less large to moderate correlations (See Table 4). One of the reasons therefore is the absence of a golden standard for assessment of movement skill development. The next paragraphs include additional information on reliability and validity of the selected tests.

The Movement-ABC test (Dutch version) complies with Evers et al.'s (2000a; 2000b) parameters on reliability. Van Waelvelde et al. (2007a) confirm reliability of the Movement-ABC for the identification of mild to moderate motor impairment in young children. However, because of the substantial standard error of measurement (SEM) of 2.4 and a learning effect that might occur, repeated testing at short time intervals and monitoring children with the Movement-ABC are not recommended (Van Waelvelde et al., 2007a). Croce et al. (2001) found high intraclass correlation coefficients (ICCs) ($r = 0.92$ to 0.98) for test reliability and concluded that stable values over a one week period were found when using the Movement-ABC. Chow and Henderson's (2003) results supported inter-rater and test-retest reliability showing ICCs all above $r = 0.95$ except for one single item. More recently, evidence was found for Movement-ABC 2 reliability in dichotomized motor classification using the 15th percentile point as cut off as results showed Kappa correlation coefficients ranging between $\kappa = 0.94$ and 1.00 (Henderson et al., 2007). Inter-rater and test retest correlation coefficients range from $r = 0.92$ to 1.00 and from $r = 0.62$ to 0.92 except for one item in the two oldest age bands. Reliability tests on three year olds resulted in moderate correlation coefficients ($r = 0.49$) (Henderson et al., 2007).

PDMS-2 reports reliability coefficients for subgroups as well as for the total sample (Follio and Fewell, 2000). Van Hartingsveldt et al. (2005) conclude that PDMS-2 (fine motor scales) has large test-retest and inter-rater reliability, the coefficients varied from $r = 0.84$ to 0.99 .

The TGDM-2 reports internal consistency and stability coefficients (Ulrich, 2000). TGDM-2 test-retest reliability, however, shows no improvement compared to the former TGDM's first edition. The procedure analyzing 30 completed protocols twice does not match a test-retest situation and cannot be considered as a measurement of stability over time (Simons and Van Hombeeck, 2003).

ICCs for test-retest reliability ranged from $r = 0.43$ to 0.93 on the individual test items and between $r = 0.61$ and 0.74 on total scores for qualitative and quantitative measures of the MMT (Kroes et al., 2004). A possible explanation for these moderate correlation coefficients was that some coefficients were process observations, which were far more difficult to score than product oriented observations. ICC for total score reliability coefficient was $r = 0.74$.

Test-retest reliability for the BOTMP has been found to be large: $r = 0.86$ to 0.89 for the Full Battery composite score (Bruininks, 1978). Additional support was found for inter-rater reliability and provides greater

confidence to use the test in children, with and without motor problems (Wilson et al., 2000). Yoon et al. (2006) conclude on the BOTMP to be a reliable tool to assess motor performance in children. These authors recommend pilot testing for feasibility testing in the target group. Hassan (2001) found small to large correlation coefficients (ranging from $r = 0.22$ to 0.80) for internal consistency in a sample of children from the United Arab Emirates when using the BOTMP SF (Short Form). This author doubts the BOTMP-SF's suitability to use in children with different types and severities of a disability. In the BOT-2 split-half methods were added to report internal consistency and test-retest and inter-rater reliability was reported for a sample of one hundred and forty three children (Peerlings, 2007).

In their test manual the authors of the MOT 4-6 concluded that neither factor analysis nor cluster analysis provided an adequate factor structure. Construct and content validity of the test items have therefore been explained based on movement skill literature. Using the KTK as reference for measuring concurrent validity a large correlation coefficient of $r = 0.68$ was found after correction for age (Zimmer and Volkamer, 1987). Kambas et al. (2002) constructed normative data tables for Greek children for MOT 4-6 use, because the authors assumed that Greek and German children might differ in motor development. Older children showed significantly better performances than younger ones (Kambas et al., 2002). The test's capacity to show developmental enhancement in healthy Greek children was confirmed (Kambas and Aggelousis, 2008).

Many studies on the usability of the Movement-ABC reported US norms to be valid in different populations, suggesting that there is no or little impact of cultural differences (Smits-Engelsman, 1998; Rosblad and Gard, 1998; Sigmundsson and Rostoft, 2003). Caution, however, is required as some authors suggest that larger studies are needed to address the question whether US norms are valid for their population (Rosblad and Gard, 1998; Sigmundsson and Rostoft, 2003) or that the validity is not supported in all ages (Livesey et al, 2007; Van Waelvelde et al., 2008). Chow et al. (2006) suggests adjustments to some test items would be desirable to be used in China as results showed both significant within and cross cultural differences in their study. Since Evers et al. (2000a; 2000b) expressed concern on the absence of criterion validity of the Movement-ABC, additional concurrent validity studies with several other well established tests, e.g. PDMS 2, BOTMP and KTK have been published (Croce et al., 2001; Van Waelvelde et al., 2004; Van Waelvelde et al., 2007b). These studies confirmed concurrent validity and supported the ability of the Movement-ABC to detect children with mild to moderate impairment. At the same time some authors stated that these tests are not interchangeable and that test choice should depend on the specific purpose of the assessment (Van Waelvelde et al., 2007b). Concurrent validity Pearson's product correlation coefficients ranged between 0.60 and 0.90 between the Movement ABC, the BOTMP-LF (Long Form) and SF (Croce et al., 2001). These authors recommend the Movement ABC to be used when short and simple assessment of a child's motor perform-

ance is required. Whether Evers (2000b) argument on the representativeness of the Movement-ABC's normative data for the Dutch population will be replicated by the new normative data collection using the Movement-ABC 2 in Dutch children, will be known when these are published. The Movement-ABC is found to be valid to monitor individual change in children with mild to moderate motor impairments (Leemrijse et al., 1999) and to detect mild to moderate impairment in preschool children (Van Waelvelde et al., 2007a). Further evaluation of gender differences for each of the test items might also be beneficial (Livesey et al., 2007).

PDMS and PDMS-2 scores showed large correlations (0.71 and 0.75), their mean scores, however, differed significantly which indicated that both versions were not equivalent in 4-year old children (Darrah et al., 2007; Tieman et al., 2005). The PDMS-2 claims representative data for the current US population. Follio and Fewell, (2000) reported support for the PDMS-2's construct validity on group differentiation for a variety of subgroups (e.g. physically handicapped) as well as for the general population. Cross cultural validity for European children, however, cannot be guaranteed. Vanvuchelen et al. (2003) emphasize the need for normative data collection in Flanders, because original normative do not permit sufficient accuracy in the detection of developmental delay. The PDMS-2 is less sensitive to mild motor impairment in a Flemish population than the Movement-ABC (Van Waelvelde et al., 2007b). In contrast with confirmation of convergent validity with the fine motor section on the Movement-ABC, Van Waelvelde et al. (2007b) found no evidence for discriminant validity between fine and gross motor sub scores. Van Hartingsveldt et al. (2005) conclude that the PDMS-2 (fine motor section) might not be sensitive enough for children with fine motor problems.

For the KTK factor analysis confirms only one factor. The KTK test has valid and separate normative data for boys and girls. A screening for cultural and geographical differences shows no significant results. Suitability of the normative data was not confirmed for some other European countries (Smits-Engelsman et al., 1998). The authors suggest that normative data should be adjusted for different populations.

Construct validity of the TGDM was supported in a larger sample of Greek children ($n = 664$, age 3 to 10) and cross generalization was shown with Alpha coefficients $\alpha = 0.75$ and 0.74 by Evaggelina et al. (2002). TGDM-2 validity is shown for a wide variety of subgroups as well as for the general population of children aged 3 to 10 years. Age norms were divided into half-year increments for both subtests (Ulrich, 2000). Construct validity has been confirmed for other populations (Wong and Cheung, 2007; Simons et al., 2004)

The MMT test was validated through one expert's (> 20 years of experience) global judgment. The aim was to determine whether qualitative aspects of motor behaviour could help distinguish between children with and without normal motor performance as judged by the school doctor. Specificity and sensitivity were calculated for different cut-off points (Kroes et al., 2004). Predictive validity for the detection of ADHD (Attention Deficit

Hyperactivity Disorder) has been observed for the qualitative aspects of the MMT, motor performance was not predictive for ODD (Oppositional Defiant Disorder) and CD (Conduct Disorder) (Kroes et al., 2002).

The BOTMP-SF (Short Form) was suggested to be used for screening purposes and therefore examined for construct validity in different populations. The BOTMP-SF was found to be valid to assess Greek pre- and primary schoolchildren (ages 5 to 8 years) (Kambas and Aggeloussis, 2006). These authors recommended caution and further study on validity is required for the use of the BOTMP-SF in children with disabilities. Venetsanou et al. (2007) conclude that the BOTMP-SF is not a valid tool to identify motor impairment in Greek 5-year-olds. Further research is needed to precise these findings. Hassan (2001) also found support for construct validity of the BOTMP-SF in 6 to 11 year-olds and concluded that the BOTMP-SF could be used to assess children's motor performance in the United Arab Emirates. According to Croce et al. (2001) the BOTMP Full Battery Composite Score correlated largely ($r = 0.76$) with the Movement-ABC. The BOTMP, however, was considered highly false negative in detecting motor delay at school age (Flegel and Kolobe, 2002). The study of Duger et al. (1999) showed that gross and fine motor skills in early childhood varied for age, sex and academic learning. Successful children had better movement performances than unsuccessful children. The outcome of this study supported validity of the BOTMP in a sample of 4 to 11 year-old children and the authors suggest that the Bruininks Oseretsky test can be useful to investigate unexplored aspects of motor development. The BOTMP's use as an assessment tool to measure motor deficiencies has been questioned in some studies (Yoon et al., 2006).

BOT-2 validity has been proven for content and factor analysis correlation coefficients support its theoretical structure (Bruininks and Bruininks, 2005). BOT-2 theoretical and empirical sources of validity evidence concern test content, item fit and clinical group differences. Content development focused on eliminating less effective BOTMP items and identifying new ones, accomplished through conducting a product survey and tree focus groups. Each new item went through three stages of development: a pilot study, a national try out and standardization. At each stage item analysis was performed and user feedback was collected. Strong relationships between the BOT-2 and other tests [PDMS-2, BOTMP, TVMS-R (Test of Visual Motor Skills – Revised (Gardner, 1995))] provide additional evidence for validity of the changes made to improve measurement.

Normative data and origin of the motor assessment tools

When using a motor development test in an educational setting, a reference is needed to rank the performance of a child. However, prudence is recommended because the normative data are often based on small samples, are rather old or aim at motor deficiencies rather than motor capacities. Table 5 shows information on the normative data, along with the authors and the origin of the tool (country), the specific age range, the number of children

Table 5. Normative data.

Test	Author	Year	Origin	Age band	n	Sample
MOT 4-6	Zimmer and Volkamer	1987	Germany	4:0 - 6:11	548	Preschool, Germany Elementary school, Germany
M-ABC (original)	Henderson, S.E. and Sugden, D.A.	1992	USA	4:0 - 12:11	1234 4-6: n = 493, 7-8: n = 264, 9-10: n = 257, 11-12+: n = 220	American children (demographic, origin and gender)
M-ABC (Dutch)	Smiths-Engelsman	1998	The Netherlands	4:0 - 12:11	549 4-6: n = 179, 7-8: n = 141, 9-10: n = 139, 11-12+: n = 70	Dutch children, Physiotherapist administrators, The Netherlands
PDMS 2	Folio and Fewell	2000	USA	0-6:11	2003	Representative of US population (1997-1998), Therapist administrators
KTK	Kiphard and Shilling	1974	Germany	5:0-14:11	1128	Elementary /high school and remedial school, Germany (1973-1974)
TGMD-2	Ulrich	2000	USA	3-10:11	1208	American children (demographic origin, race and gender)
MMT	Vles, Kroes and Feron	2004	The Netherlands	5-6	800	Regular elementary schools, The Netherlands
BOT-2	Bruininks and Bruininks	2005	USA	4-21	1520 4-6: n = 510	US population of 2004 - 2005 (sex, race/ethnicity, socioeconomic status, and disability status)

in the sample and sampling method. Most motor developmental assessment tools in this review have normative data that are only representative for the US population. Some of the data differentiates for ethnicity, race, gender, etc. Only the Movement-ABC test has relatively recent normative data for European children. The KTK and MOT 4-6 and the MMT tests also provide normative data for European children. Some of these data have as a minus point that they are rather dated and limited because they were mainly used in the country of origin (See Table 5).

Critical considerations on the use of motor assessment tools

The most fundamental criticism on movement skill assessment tools is that they do not have the same psychometric quality as tools used to assess cognitive development (Netelenbos, 2001a; 2001b). According to this author, there are five main reasons for this particular shortcoming. First, cognitive development is considered the most important developmental goal and since there is limited evidence that information on movement skill development supports a better understanding of cognitive development, the interest in movement skill development is limited. Secondly, PE is often not valued as high as other subjects. Third, there is no evidence for the existence of undivided motor capacity. Measuring a large number of items using simple tasks might be a possible solution, but will become too time consuming. Fourth, possibly contradictory results on gender differences do not stimulate the creation of gender neutral, reliable and valid assessment tools. According to Netelenbos (2001b)

gender differences emerge at eleven or later, but not all authors agree with this statement. For example; Pendersen et al. (2003) and Ulrich (2000) report on clear gender differences for gross and fine movement skills and Van Waelvelde et al. (2003) state it is a shortcoming of many movement tests for children not to offer separate norms for boys and girls. Finally, there are great discrepancies between children of the same age range. Especially when total test scores are used for analysis, a test user should be aware of possible low correlation between different motor tasks. There are no specific age norms for the acquisition of FMS. The complexity of movement skills assessment reflects a multifactor identity of the motor system, the possible presence of gender or cultural differences and the large variance within children of the same age. In a diagnosis process the use of more than one assessment tool is recommended (Netelenbos, 2001b).

Table 6 shows strengths and weaknesses of the different movement skill assessment tools for use with typically developing children in a preschool research setting.

Many different factors influence the choice of a movement assessment tool. To make a selection of the test(s) that will be used in educational research settings, the following criteria should be considered:

- purpose of assessment: general motor proficiency, fine motor or gross motor proficiency assessment, prevalence of motor impairment;
- age specificity and appropriateness of the test;
- simplicity of the test (instruction and demonstration should be short and simple);
- training easiness of examiners and observers;
- cultural similarity between norm and test group;
- proportion of tested items in relation to test time.

Table 6. Strengths and weaknesses of the assessment tools for use in educational research for preschool children.

Motor assessment tool	Strengths	Weaknesses
MOT 4-6	Age appropriate for preschool children Appropriate for use in an educational setting Short and clear score sheet Highly efficient (items tested / assessment time) Score refers to general fundamental movement skill performance Provides information on skill mastering: both beneath and above skill level	Older normative data No recent revision of the test available Little information on the test and test results available in international literature
Movement ABC	Age appropriate for preschool children International normative data is available Quantitative and qualitative assessment Highly suitable for impairment detection Usable in educational setting Large amount of studies on the psychometric qualities of the test	No information of skill mastery above skill level Rather low efficiency (unfavorable proportion of items tested versus assessment time), Not specifically designed for young children
PDMS 2	Age appropriate for preschool children Very detailed assessment instrument Subtest composites can be assessed separately Provides information on skill mastering: both beneath and above skill level Inclusion of Qualitative aspects of movement behavior	Specifically designed to detect deficits / motor impairment No short form available Taking of the complete test is rather long for young children Absence of normative data on European children
KTK	Quick screening of stability skill Still considered as very reliable and highly valued for its accuracy and standardization (Simons, 2004).	Older normative data Less age appropriate for the assessment of preschool children One-sided information on movement skill development (balance)
TGMD-2	Age appropriate for preschool children Inclusion of qualitative aspects of movement behavior Emphasis on object control movement skill development Provides information on skill mastering: both beneath and above skill level	Does not evaluate fine nor stability movement skill development For Europe the battery is not free from cultural differences
MMT	Combination of quantitative and qualitative assessment items Highly efficient (items tested / assessment time) Provides information on skill mastering: both beneath and above skill level	Developed for a very small age range Especially designed for the detection of ADHD
BOT-2	Age appropriate for preschool children Very detailed assessment instrument Subtest composites can be assessed separately Short Form assessment for general movement skill development available Provides information on skill mastering: both beneath and above skill level Inclusion of qualitative aspects of movement behavior Large amount of evidence on psychometric qualities of the test	Emphasis lies on detection of deficits Absence of European normative data Complicated to receive a test kit Standardization of examination is rather difficult Taking of the complete test is rather long for young children A large test room is needed for the running item Confusing score sheet

The overview of normative data underlines the one-sided normative US samples used in assessment tools and the scarcity of validity for cross cultural use of the tests (See Table 5). The available data are either dated and/or based on small age group samples. This supports the importance of including European preschool children to provide adequate normative data for cross cultural use of these tests. The challenges involve choosing appropriate test items because clearly identified differences between movement skill development of American and European children exist and as already mentioned, there is a shortage of up to date information on movement skill development and performance in Europe (Peerlings, 2007; Simons and Van Hombecq, 2003; Vanvuchelen et al., 2003).

In a research context recommendations for test use will depend on the purpose of the study. When the purpose is the assessment of preschool children's general capabilities, the use of the MOT 4-6 is recommended. When the focus is specifically on stability skills, the KTK test would be more appropriate. Further research and wider use of the MMT might reveal the potential to be used in an educational research setting. When the purpose, however, is assessing the prevalence of motor im-

pairment among preschool children the Movement-ABC would be suitable to use. Nevertheless, the sensitivity and specificity of the Movement ABC-2 should be further investigated. More complex instruments such as the BOT-2 and PDMS-2, are more appropriate to assess smaller groups of children. These tools have a rather time consuming nature. Separate scores for fine and gross movement performance are obtained using the PDMS-2 allowing relative differences in gross and fine movement performance of children between birth and 6 years of age (Tieman et al., 2005). In agreement with Wilson et al. (1995) it can be concluded that BOTMP and its successor can be considered useful when changes in performance are evaluated. The TGDM-2 needs adaptation to fit in a European context because the test is too culturally dependent (e.g. the object control subtest: striking a stationary ball and overhand throwing).

A general remark on the normative samples used in the tests is that only small samples for each age group of developing children were used. At this age children's FMS develop relatively fast (Netelenbos, 2001a) and age categories of 6 months should be preferred to 2 year categories.

Conclusion

In conclusion, the primary goal of most of the reviewed instruments is to detect deficiencies in movement skill development. Most of the studies using these tools hardly discuss the variation in motor skill development of typically developing children and most of the data on typically developing children have been gathered by professionals educated for the detection of irregular motor behavior. A suggestion for further research is that PE teachers are also involved in normative data collection.

However, the goal of early detection can only be pursued if there is a precise and up-to-date description of typical movement skill development and performance of a particular population. Therefore, it is just as important that long term follow up of FMS development is performed and continued, and that well considered measures are taken to enhance FMS development.

Acknowledgments

This work was supported by the Vrije Universiteit Brussel [OZR989BOF].

References

- Barnett, A. and Peters, J. (2004) Motor proficiency assessment batteries. In: *Developmental motor disorders: A neuropsychological perspective*. Eds: Dewey and Tupper. New-York: Guilford, 67-109.
- Bertrands, E., Medts De, C. and Descheppere, G. (2003) *Kleuterstappen in beweging, bewegingszorg voor het jonge kind*, Leuven: Acco.
- Berk, L. (2003). *Child Development* (sixth edition ed.). Boston: Allyn & Bacon.
- Bös, K. (2003) *Motorische Leistungsfähigkeit van Kindern und Jugendlichen*. Erster Deutscher Kinder- und Jugendsportbericht. Schorndorf: Verlag Karl Hofman. 1-23.
- Bruininks, R.H. and Bruininks, B.D. (2005) *Test of Motor Proficiency*. 2nd edition. Manual.: AGS Publishing, Circle Pines.
- Bruininks, R. H. (1978) *Bruininks Oseretsky test of motor proficiency*. American Guidance Service, Circle pines-Minnesota.
- Burton, A.W. and Miller, D.E. (1998) *Movement skill assessment*. Human Kinetics, Champaign.
- Chow, S. and Henderson, S., (2003) Interrater and test-retest reliability of the Movement Assessment Battery for Chinese preschool children. *American Journal of Occupational Therapy* **57**(5), 574-577.
- Chow, S., Barnett, A.L. and Henderson, S.E. (2001) The Movement Assessment Battery for Children: A comparison of 4-year-old to 6-year-old children from Hong Kong and the United States. *American Journal of Occupational Therapy* **55**, 55-61.
- Chow, S., Hsu, Y., Henderson, S.E., Barnett, A. and Kai, Lo. S. (2006) The Movement ABC: a cross-cultural comparison of pre-school children from Hong Kong, Taiwan, and the USA. *Adapted Physical Activity Quarterly* **23**, 31-48.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. 2nd edition. New Jersey: Lawrence Erlbaum.
- Croce, R.V., Horvat, M. and McCarthy, E. (2001) Reliability and concurrent validity of the Movement Assessment Battery for Children. *Perceptual and motor skills* **93**, 275-280.
- Darrah, J., Magill-Evans, J., Volden, J., Hodge, M. and Kembhavi, G. (2007) Scores of typically developing children on the Peabody Developmental Motor Scales: infancy to preschool. *Physical and Occupational Therapy in Pediatrics* **27**(3), 5-19.
- Davies, M. (2003) *Movement and dance in early childhood, second edition*. Paul Chapman Publishing, London.
- Duger, T., Bumin, G., Uyanik, M., Aki, E. and Kayihan, H. (1999). The assessment of Bruininks-Oseretsky test of motor proficiency in children. *Pediatric Rehabilitation* **3**(3), 125-131.
- Eurydice (2002) *Key Data on Education in Europe 2002 / European Commission, Eurydice, Eurostat*. Office for Official Publications of the European Communities, Luxembourg.
- Evangelinou, C., Tsigilis, N. and Papa, A. (2002) Construct validity of the Test of Gross Motor Development: a cross-validation approach. *Adapted Physical Activity Quarterly* **19**, 483-495.
- Evers, A., van Vliet-Mulder and Groot, C.J. (2000a) *Documentatie van tests en testresearch in Nederland. Deel I: Testbeschrijvingen*. Van Gorcum and Comp. B.V., Assen.
- Evers, A., van Vliet-Mulder and Groot, C.J. (2000b) *Documentatie van tests en testresearch in Nederland. Deel II: Testresearch*. Van Gorcum and Comp. B.V., Assen
- Flegel, J. and Kolobe, T.H.A. (2002) Predictive validity of the Test of Infant Motor Performance as measured by the Bruininks-Oseretsky Test of Motor Proficiency at school age. *Physical Therapy* **82**(8), 762-771.
- Folio, M.R. and Fewell, R.R. (1983) *Peabody Developmental Motor Scales and Activity Cards*. DLM teachers resources, Allen-Texas.
- Folio, M.R. and Fewell, R.R. (2000) *Peabody Developmental Motor Scales. Examiners manual*. Pro-ED. Inc., Austin-Texas.
- Gallahue, D.L. and Ozmun, J.C. (2006) *Understanding Motor Development. Infants, Children, Adolescents, Adults*. Sixth Edition. McGraw-Hill, New York.
- Gallahue, D. and Donnelly, F. (2003) *Developmental physical education for all children*. Human Kinetics, Champaign.
- Gardner, M.F. (1995) *Test of Visual-Motor Skills-Revised*. Psychological and educational Publications Inc., Hydesville, CA.
- Gheysen, F., Loots, G. and Van Waelvelde, H. (2008) Motor development of deaf children with and without cochlear implants. *The Journal of Deaf Studies and Deaf Education* **13**(2), 215-224.
- Hassan, M.M. (2001) Validity and reliability for the Bruininks-Oseretsky-Test of Motor Proficiency-short form as applied in the United Arab Emirates culture. *Perceptual and Motor Skills* **92**, 157-166.
- Haywood K.M. and Getchell N. (2005) *Life span motor development*. 4th edition. Human Kinetics, Champaign.
- Henderson, S.E. and Sugden, D.A. (1992) *Movement Assessment Battery for children*. Sidcup. Therapy skill builders, Kent-England.
- Henderson, S.E., Sugden, D.A. and Barnett, A.L. (2007) *Movement Assessment Battery for children – 2 Examiner's Manual*. Harcourt Assessment, London.
- Kambas, A. and Aggeloussis, N. (2006) Construct validity of the Bruininks-Oseretsky test of Motor Proficiency-short form for a sample of Greek preschool and primary school children. *Perceptual and Motor Skills* **102**, 65-72.
- Kambas, A. and Aggelousis N. (2008) Erste Ergebnisse mit dem MOT 4-6 im griechischen Kindergarten Eine Voruntersuchung. Unpublished manuscript, University of Thrace, Greece.
- Kambas, A., Aggelousis, N., Proviadaki, I., Taxildaris, K., Mavromatis, G. (2002) Standardization of the "MOT 4-6" motor development assessment battery in Greek preschool children: a pilot study. *Exercise and Society Journal of Sport Science* **30**, 28-37.
- Kiphard, E.J. and Shilling, F. (1974) *Körperkoordinationstest für Kinder*. Beltz test, Weinheim.
- Kiphard, E.J. and Shilling, F. (2007) *Körperkoordinationstest für Kinder 2, überarbeitete und ergänzte Auflage*. Beltz test, Weinheim.
- Kroes, M., Kessels, A.G.H., Kalf, A.C., Feron, F.J.M., Vissers, Y.L.J., Jolles, J. and Vles, J.S.H. (2002). Quality of movement as predictor of ADHD: results from a prospective population study in 5- and 6-year-old children. *Developmental Medicine and Child Neurology* **44**, 753-760.
- Kroes, M., Vissers, Y.L.J., Sleijpen, F.A.M., Feron, F.J.M., Kessels, A.G.H., Bakker, E., Kalf, A.C., Hendriksen J.G.M., Troost, J., Jolles, J. and Vles, J.H.S. (2004) Reliability and validity of a qualitative and quantitative motor test for 5- to 6-year-old children. *European Journal of Paediatric Neurology* **8**, 135-143.
- Leemrijse, C., Meijer, O.G., Vermeer, A., Lambregts, B. and Adèr, H.J. (1999) Detecting individual change in children with mild to moderate impairment: the standard error of measurement of the Movement ABC. *Clinical Rehabilitation* **13**, 420-429.
- Livesey, D., Coleman, R. and Piek, J. (2007) Performance on the Movement Assessment Battery for Children by Australian 3- to 5-year-old children. *Child: Care, Health and Development* **33**(6), 713-719.
- Netelenbos, J.B. (2001a) *Motorische ontwikkeling van kinderen, handboek 1, introductie*. Boom, Amsterdam.
- Netelenbos, J.B. (2001b) *Motorische ontwikkeling van kinderen, handboek 2, theorie*. Boom, Amsterdam.
- Pedersen, A.V., Sigmondsson, H. and Whiting, H.T.A. (2003) Sex differences in lateralisation of fine manual skills in children. *Experimental Brain Research* **149**(2), 249-251.

- Peerlings, W. (2007) *Bruininks-Oseretsky Test of Motor Proficiency*. 2nd edition. (BOT-2). In *Actuele Themata uit de psychomotorische therapie*, Johan Simons (ed.). Acco, Leuven/Leusden.
- Petermann, F. (Hrsg.). (2008) *Movement Assessment Battery for Children-2 (Movement ABC-2)*. Frankfurt: Pearson PLC.
- Rösblad, B. and Gard, L. (1998) The assessment of children with developmental coordination disorders in Sweden: A Preliminary Investigation of the Suitability of the Movement ABC. *Human Movement Science* **17**, 711-719.
- Rosenbaum, P., Missiuna, C. and Johnson, K. (2004) *Longitudinal Assessment of Motor Development in Epidemiologic Research for the National Children's Study*. Report for the NCS by Battelle Memorial Institute.
- SIG (2005) *Kleuters veilig oversteken: kleutervolgsysteem*. SIG, Destelbergen.
- Sigmundsson, H. and Rostoft, M. (2003) Motor Development: exploring the motor competence of 4-year-old Norwegian Children. *Scandinavian Journal of Educational Research* **47(4)**, 451-459.
- Simons, J. (2004) *Introductie tot de psychomotoriek*. Garant, Antwerpen-Appeldoorn.
- Simons, J., Daly, D., Theodorou, F., Caron, C., Simons, J. and Andoniadou, E. (2008) Validity and reliability of the TGMD-2 in 7-10-year-old Flemish children with intellectual disability *Adapted Physical Activity Quarterly* **25(1)**, 71-82.
- Simons, J. and Van Hombeek, C. (2003) Toepasbaarheid van de test of Gross Motor Development 2nd ed. *Kinevaria* **39(4)**, 16-21.
- Smits-Engelsman B.C.M., Henderson, S.E. and Michels, C.G.J. (1998) The assessment of children with Developmental Coordination Disorders in the Netherlands: The relationship between the Movement Assessment Battery for Children and the Körperkoordinations Test für Kinder. *Human Movement Science* **17**, 699-709
- Smits-Engelsman, B.C.M. (1998) *Movement Assessment Battery for children, Handleiding*. Swets & Zeitlinger B.V., Lisse.
- Task Force on Standards for Measurement in Physical Therapy (1991) Standards for tests and measurements in Physical Therapy practice. *Physical Therapy* **71**, 589-622.
- Tieman, B. L., Palisano, R. J. and Sutlive, A. C. (2005). Assessment of motor development and function in preschool children. *Mental retardation and Ddevelopmental Disabilities Research Reviews* **11(3)**, 189-196.
- Tripathi, R., Joshua, A.M., Kotian, M.S. and Tedla, J.S. (2008) Normal motor development of Indian children on Peabody Developmental Motor Scales-2 (PDMS-2). *Pediatric Physical Therapy* **20(2)**, 167-172.
- Soppelsa R. and Albaret J.-M. (2004) *Manuel de la Batterie d'Evaluation du Mouvement chez l'Enfant*. Paris : ECPA.
- Ulrich, D.A. (1985) *Test of Gross Motor Development*. Pro-ED. Inc., Austin, Texas.
- Ulrich, D.A. (2000) *Test of Gross Motor Development, 2nd ed. Examiner's manual*. Pro-ED. Inc., Austin, Texas.
- Vallaey, M. and Vandroemme, G. (1999) *Psychomotoriek bij kinderen*. Leuven: Acco.
- van Gelder W jr. and Stroes H, 2002, *Leerlingenvolgsysteem spelen en bewegen, over observeren, registreren en extra zorg*. Elsevier, Maarssen.
- Van Hartingsveldt, M.J., Cup, E.H.C. and Oostendorp, R.A.B. (2005) Reliability and validity of the fine motor scale of the Peabody Developmental Motor Scales-2. *Occupational Therapy International* **12(1)**, 1-13.
- Van Waelvelde, H., De Weerd, W., De Cock, P. and Smits Engelsman B.C.M. (2003) Ball catching. Can it be measured? *Physiotherapy Theory and Practice* **19**, 259-267.
- Van Waelvelde, H., De Weerd, W., De Cock, P., Smits-Engelsman B.C.M. (2004) Aspects of the validity of the Movement Assessment Battery for Children. *Human Movement Science* **23**, 49-60.
- Van Waelvelde, H., Peersman, W. Lenoir, M. and Smits Engelsman B.C.M. (2007a) The reliability of the Movement Assessment Battery for Children with mild to moderate motor impairment. *Clinical Rehabilitation* **21**, 465-470.
- Van Waelvelde, H., Peersman, W. Lenoir, M. and Smits Engelsman B.C.M. (2007b) Convergent validity between two motor tests: Movement-ABC and PDMS-2. *Adapted Physical Activity Quarterly* **24**, 59-69.
- Van Waelvelde, H., Peersman, W., Lenoir, M., Smits Engelsman, B.C.M., Henderson, S.E. (2008) The movement assessment battery for children: Similarities and differences between 4- and 5-year-old children from Flanders and the United States. *Pediatric Physical Therapy* **20(1)**, 30-38.
- Vanvuchelen, M. Mulders, H. and Smeyers, K. (2003) Onderzoek naar de bruikbaarheid van de recente Amerikaanse Peabody Developmental Scales-2 voor vijfjarige, Vlaamse kinderen. *Signaal* **45**, 24-41.
- Venetsanou, F., Kambas, A., Aggeloussis, N., Serbezis, V. and Taxildaris, K. (2007) Use of the Bruininks-Osetetsky Test of Motor Proficiency for identifying children with motor impairment. *Developmental Medicine & Child Neurology* **49(11)**, 846-848.
- Vles J.S.H., Kroes M. and Feron F.J.M. (2004) *MMT: Maastrichtse Motoriek Test*. Pits BV, Leiden.
- Wiaert, L. and Darrah, J. (2001) Review of four tests of gross motor development. *Developmental Medicine and Child Neurology* **43**, 279-285.
- Wilson, B.N., Polatajko, H.J., Kaplan, B.J. and Faris, P.D. (1995) Use of the Bruininks-Oseretsky Test of Motor Proficiency in occupational therapy. *American Journal of Occupational Therapy* **49**, 8-17.
- Wilson, B.N., Kaplan, B.J., Crawford, S.G. and Dewey, D. (2000) Interrater reliability of the Bruininks-Oseretsky Test of Motor Proficiency - Long Form. *Adapted Physical Activity Quarterly* **17**, 95-110.
- Wong, K.Y.A. and Cheung, S.Y. (2007) Structural Validity of the Test of Gross Motor Development-2. *AAHPERD National Convention and Exposition, March 13 - 17, 2007* Baltimore, MD. (Poster presentation)
- Yoon, D, Scott, K. and Hill, M. (2006) Review of three tests of motor proficiency in children. *Perceptual & Motor Skills* **102**, 543-551.
- Zimmer, R. and Volkamer, M. (1987) *Motoriktest für vier- bis sechsjährige Kinder (manual)*. Beltztest, Weinheim.
- Zimmer, R. (2006) Diagnostik der Bewegungsentwicklung von Kindern: Beobachten – Einschätzen – Dokumentieren. *2nd international conference Learning processes in early childhood and assessment issues, June, 22-24, Bolzano-Italy*. Book of abstracts. 15.

Key points

- This review discusses seven movement skill assessment tool's test content, reliability, validity and normative samples.
- The seven assessment tools all showed to be of great value. Strengths and weaknesses indicate that test choice will depend on specific purpose of test use.
- Further data collection should also include larger data samples of able bodied preschool children.
- Admitting PE specialists in assessment of fundamental movement skill performance among preschool children is recommended.
- The assessment tool's normative data samples would benefit from frequent movement skill performance follow-up of today's children.

Abbreviations

MOT 4-6: Motoriktest für vier- bis sechsjährige Kinder,
 M-ABC: Movement Assessment Battery for Children,
 PDMS: Peabody Development Scales,
 KTK: Körper-Koordinationstest für Kinder,
 TGDM: Test of Gross Motor Development,
 MMT: Maastrichtse Motoriektest,
 BOTMP: Bruininks-Oseretsky Test of Motor Proficiency.
 ICC: intraclass correlation coefficient,
 NR: not reported,
 GM: gross motor,

LV: long version,
 SV: short version,
 LF: long form,
 SF: short form,
 STV: subtest version,
 SEMs standard errors of measurement,
 TMQ: Total Motor Quotient,
 TMC: Total Motor Composite,
 CSSA: Comprehensive Scales of Student Abilities
 MSEL: Mullen Scales of Early learning: AGS Edition
 AUC: Areas under curve
 BC: Battery composite
 ROC: Receiver operating characteristic

AUTHORS BIOGRAPHY



Wouter COOLS

Employment

Research assistant, PhD-student at the Department of Movement Education and Sport Training, Faculty of Physical Education and Physiotherapy.

Degree

MSc

Research interests

Physical education with focus on physical activity and movement skill development.

E-mail: wcools@vub.ac.be



Kristine De MARTELAER

Employment

Professor at the Department of Movement Education and Sport Training, Faculty of Physical Education and Physiotherapy.

Degree

PhD

Research interests

Pedagogy with focus on experiences and expectations of children and teachers / coaches with PE, sport and BLS, competences and job profile of teachers PE and youth coaches, Physical Education Teacher Education, motor development of young children in relation with the movement culture and education through dance.

E-mail: kdmartel@vub.ac.be



Christiane SAMAHEY

Employment

Department of Movement Education and Sport Training, Faculty of Physical Education and Physiotherapy, Vrije Universiteit Brussel.

Degree

MSc

Research interests

Psychomotor development and behavior in the age range of 0 to 8years old.

E-mail: Christiane.Samaey@vub.ac.be



Caroline ANDRIES

Employment

Professor at the Vrije Universiteit Brussel, Faculty of Psychology and Educational Sciences, Head of the Department of Developmental and Lifespan Psychology.

Degree

PhD

Research interests

Developmental psychology with a focus on epidemiology and prevention of risk behavior in adolescents and dyslexia.

E-mail: candries@vub.ac.be

✉ MSc Wouter Cools

Vrije Universiteit Brussel, Faculty of Physical Education and Physiotherapy (Faculty LK), Department of Movement Education and Sport Training (BETR), Pleinlaan 2 / L209, B-1050 Brussel, Belgium