Agreement Between the KTK3+ Test and the Athletic Skills Track for Classifying the Fundamental Movement Skills Proficiency of 6- to 12-Year-Old Children

M. de Niet[1]*, S. W. J. Platvoet[1], J. J. A. A. M. Hoeboer[2], A. M. H. de Witte[2], S. I. de Vries[2] and J. Pion[1,3]

[1]School of Sport and Exercise, Talent Identification and Talent Development in Sport, HAN University of Applied Sciences, Nijmegen, Netherlands, [2]Centre of Expertise Health Innovation, Research Group Healthy Lifestyle in a Supporting Environment, The Hague University of Applied Sciences, The Hague, Netherlands, [3]Department of Movement and Sports Sciences, Faculty of Medicine and Health Sciences, Ghent University, Ghent, Belgium

The main aim of this study was to determine the agreement in classification between the modified KörperKoordinations Test für Kinder (KTK3+) and the Athletic Skills Track (AST) for measuring fundamental movement skill levels (FMS) in 6- to 12-year old children. 3,107 Dutch children (of which 1,625 are girls) between 6 and 12 years of age (9.1 ± 1.8 years) were tested with the KTK3+ and the AST. The KTK3+ consists of three items from the KTK and the Faber hand-eye coordination test. Raw scores from each subtest were transformed into percentile scores based on all the data of each grade. The AST is an obstacle course consisting of 5 (grades 3 till 5, 6–9 years) or 7 (grades 6 till 8, 9–12 years) concatenated FMS that should be performed as quickly as possible. The outcome measure is the time needed to complete the track. A significant bivariate Pearson correlation coefficient of 0.51 was found between the percentile sum score of the KTK3+ and the time to complete the AST, indicating that both tests measure a similar construct to some extent. Based on their scores, children were classified into one of five categories: <5, 5–15, 16–85, 86–95 or >95%. Cross tabs revealed an agreement of 58.8% with a Kappa value of 0.15 between both tests. Less than 1% of the children were classified more than two categories higher or lower. The moderate correlation between the KTK3+ and the AST and the low classification agreement into five categories of FMS stress the importance to further investigate the test choice and the measurement properties (i.e., validity and reliability) of both tools. PE teachers needs to be aware of the context in which the test will be conducted, know which construct of motor competence they want to measure and know what the purpose of testing is (e.g., screening or monitoring). Based on these considerations, the most appropriate assessment tool can be chosen.

Keywords: primary school, validity, measurement, motor competence, motor skill, fundamental movement skills

INTRODUCTION

In the Netherlands, as well as many other countries, the main goal of physical education in primary schools is to develop the motor competence of every child between ages 4–12. That is a huge challenge, especially as there are large differences in motor competences between children. Awareness of the importance of testing in educational setting have increased in the Netherlands among PE teachers as well as policy (makers). An easy applicable test for monitoring large groups of children over time could therefore be highly valuable to PE to better meet each child demands and could create a better understanding of motor learning and development (Cliff et al., 2009; Lloyd et al., 2014; Scheuer, 2019). To optimize the usefulness of testing for all children, the test should be able to identify the full range of performance from motor disorders to the highest proficiency levels (Platvoet et al., 2018), although PE is often focused more on motor disorders (Scheuer, 2019).

Recently, Barnett and colleagues (2016) stated the importance of fundamental movement skills (FMS) in PE. The FMS are considered to be the foundation for motor competence (Barnett et al., 2016). In daily life, FMS are essential requirements for physical activities (Henderson et al., 2008; Stodden et al., 2008; Lubans et al., 2010; Logan et al., 2018). These generic skills are the building blocks for more specific (sports) skills learned at later developmental stages (Clark, 2007; Gallahue, 2012; Lloyd et al., 2014; Loprinzi et al., 2015; Cattuzzo et al., 2016; Logan et al., 2018). FMS, can be divided in three categories (balancing, locomotor and object control skills) which include (stick) balancing, hopping, running, and throwing and catching (Gallahue, 2012). These skills have subsequently been described as learned movement patterns built up by gross motor coordination skills. The mastery of these skills should begin in early childhood. In this stage, there is an acceleration of the cortex maturation and the cortex become more organized (Gallahue and Donnely, 2003). Therefore, it is important to support PE teachers, and more in general all professionals working with children, with feasible assessment tools to help them gain insight into the actual development of children's FMS proficiency (Cools et al., 2009; Platvoet et al., 2018; Morley et al., 2019). In addition, PE teachers (and professionals) can better support children in their learning processes if they have more insight into children's FMS and are more involved in assessing those skills. With feasible assessment tools, PE teachers will be better able to meet each child demands which is required to develop the FMS proficiency of all children. Not only of those who score average, but also of those with very poor proficiency and those who are highly skilled at a young age.

In the last 2 decades, several tools to assess children's FMS have been discussed in the literature (Wiart and Darrah, 2001; Barnett and Peters, 2004; Tieman et al., 2005; Yoon et al., 2006; Cools et al., 2009; Scheuer et al., 2019; Hulteen et al., 2020). Initially, those tools were mainly developed for research and/or clinical purposes, especially to identify children with motor disabilities. Many of these tools are less suitable for use in physical education (PE) because assessment requires too much time and/or too much specific equipment. In addition to these practical limitations, many of these tests—e.g., the Test of Gross Motor Development (Ulrich, 1985; Ulrich, 2000) and the Bruininks–Oseretsky Test for Motor Proficiency 2 (Bruininks and Bruininks, 2005)—often show ceiling effects for better performers. Ceiling effects make these tests less suitable for assessing FMS proficiency in typically developing children as differentiation between (above) average performers cannot be done (Yoon et al., 2006; Cools et al., 2009). More recently, however, there has been a shift toward monitoring and evaluating typically developing children in applied settings, especially in PE. In applied settings it is important that the test is feasible: easy to use, limited in time, space, and equipment, low in costs, limited in training and qualification, and easy to administer (Cools et al., 2009; Lloyd et al., 2014; Klingberg et al., 2018). Therefore, more recent studies have focused on developing and validating tests that can be used in these applied settings (e.g., Hoeboer et al., 2016; Hoeboer et al., 2018; Herrmann and Seeling, 2017; Longmuir et al., 2017; Novak et al., 2017; Platvoet et al., 2018; Morley et al., 2019). However, the high feasibility demands also results in restrictions as for example the number of skills tested and the quality of the assessor.

Two product-oriented FMS tests are currently popular in use within PE in the Netherlands. First, the modified KörperKoordinations Test für Kinder (KTK3+) and second, the Athletic Skills Track (AST). These tests have been (re) developed in collaboration with practitioners (e.g., PE teachers), both tests have good feasibility and aim to provide these professionals with an indication of children's FMS proficiency. The reliability of the original KTK was shown by high test-retest reliability scores, $r = 0.97$ (Kiphard and Schilling, 1974), and in a three-year follow-up study in which high correlations coefficients ($r = 0.80$) were found (Vandorpe et al., 2012). The explained variance on the total KTK scores by the four subtests was high (ranges between 81 and 98%) and the content structure was also shown by a factor analysis where all test items load on the same factor (Kiphard and Schilling, 1974). Recently, the original KTK has been modified for use in applied settings. The hopping test-item was left out previously (Novak et al., 2017) and an object control task was implemented to assess skills in all three categories of FMS, resulting into the KTK-3+EHC (KTK3+) (Platvoet et al., 2018). The KTK3+ therefore consists of three items from the KTK (Kiphard and Schilling, 1974; Kiphard and Schilling, 2007) and the Faber eye-hand coordination test (Faber et al., 2014). The four isolated test items are: 1) walking backward, 2) moving sideways, 3) jumping sideways and 4) the Faber hand-eye coordination test. Platvoet and colleagues (2018) showed that the KTK3+ test measures a broad performance spectrum and differentiate between ages, which proves the concurrent validity. Together, the four items measure all three categories of the FMS i.e. balance, locomotion and object control (Platvoet et al., 2018). The Athletic Skills Track was also developed to evaluate the FMS of children aged 4–12 years in a PE setting (Hoeboer et al., 2016; Hoeboer et al., 2018). Three age-related tracks have been validated in 4- to 12-year-old children ($n = 930$) (Hoeboer et al., 2018). Depending on age, the track consists of five to seven concatenated FMS (balancing, traveling jumps, hopscotch, alligator crawl,

slaloming/running, roll and clambering) that should be completed as quickly as possible, in which individual skills were not assessed separately.

A previous study found moderate to high correlation coefficients between the KTK and AST among 4- to 12-year-old children with a range between $r = -0.60$ and $r = -0.75$ (Hoeboer et al., 2018). This study used the original KTK test, including the hopping over an obstacle subtest opposed to the KTK3+ in this study. Additionally, professionals may use these tests to classify children as well, it is therefore also relevant to determine the classification agreements between both tests, rather than just the association, as was suggested by Fransen and colleagues (2014). A high agreement score would suggest that the choice of tests has only limited influence on the classification of children based on their FMS proficiency, while a low classification agreement would suggest that the tests may assess different aspects of the same construct. Some evidence for agreement between the KTK and the AST in classifying children according to their level of FMS has been available (Hoeboer et al., 2016). That study found that the time scores for completing the AST were, on average, significantly different between the five KTK categories of high, good or normal motor giftedness or moderate or serious motor disorder. However, the study did not investigate whether individuals were classified identically, as the percentage of agreement between the tests was not calculated.

Additionally, the KTK3+ is a modified version of the original KTK. Therefore, the first aim of this study was to (re)evaluate the correlation between the (modified) tests and to investigate whether these results were comparable to previous results. The association between the AST and the KTK3 (without an object control task) was also investigated, as the AST does not contain an object control task, which may affect the association. However, the main aim of this study was to determine the agreement in classification between the KTK3+ and the AST for measuring FMS in 6- to 12-year-olds. Since both tests were developed to measure children's FMS proficiency and the AST has been validated previously by the KTK (Hoeboer et al., 2016, Hoeboer et al., 2018), we hypothesize that there is a fair level of agreement between the tests, although different skills were assessed.

MATERIALS AND METHODS

Participants

A total of 3,803 children (1,833 boys, and 1,970 girls between 6 and 12 years) in grades 3 through 8 from 21 primary schools in the Netherlands participated in this study. The schools were recruited in neighborhoods with different socio-economic statuses in two average-sized cities (150,000–170,000 inhabitants, 11 schools) and one large city (515,000 habitants, ten schools). All schools had a PE teacher. Children who were unable to complete the assessment due to physical limitations were excluded. The study procedure was approved by the ethical advisory board of the Faculty of Health at HAN University of Applied Sciences (reference number EACO 108.06/18). Schools

informed the parents prior to the testing and parents were given the opportunity to communicate with the head of the school if they did not want their child(ren) to participate. All data were anonymously recorded and processed in a secured database.

Procedures

The children were assessed with the KTK3+ (Platvoet et al., 2018) and the AST (Hoeboer et al., 2016, 2018) in October or November 2018 and performed the KTK3+ and the AST in one regular PE lesson. The test supervisors were undergraduate PE students in their last year of education. They were trained in assessing both tests to guarantee that the test protocols were used in a standardized way. Training consisted of theoretical background, protocol explanation, and training of assessing the test (approximately 4 h session). During each assessment of one class, a researcher (graduated in PE) with experience in assessing these tests supervised all assessments. Test scores on every separate attempt were scored live and directly entered in a digital database by the test supervisors through a web-based application (Sportamundi, Belgium). The web-based application also has a built-in timer and stopwatch specifically set for each test separately.

The KTK3+ consists of four subtests that measures children's gross motor coordination: KTK walking backward (3 attempts on three different sized beams, total of nine attempts, score is total number of steps), KTK moving sideways (2 attempts of 20 s, score is total number of moves), KTK jumping sideways (2 attempts of 15 s, score is total number of jumps) and hand-eye coordination (2 attempts of 30 s, score is total number of catches). For a detailed description of the KTK3+, we refer to Platvoet and colleagues (2018). The outcome measures of the four test items are the total number of correct performances within the time provided for each subtest.

The AST consists of an age-related track with several concatenated FMS that should be performed as quickly as possible. In this study, children from grades 3 through 5 completed the AST-2, which consists of the following activities: balancing (forward), traveling jumps, hopscotch, alligator crawl (backward), running (backward), pencil roll and clambering. Children from grades 6 through 8 completed the AST-3, which consists of: balancing (backward), traveling jumps, hopscotch (backward), alligator crawl (backward), slaloming (backward), forward roll and clambering. The outcome measure is the time in seconds needed to complete the track. Children had two practice runs, and one attempt that was timed in seconds and stored directly into the database. For a detailed description of the AST, see the paper by Hoeboer and colleagues (2018).

The test supervisor explained the tests to all the children. After the instructions, children were distributed at random among the tests. Before performing each test, children were asked whether they understood what they had to do. If a child replied that they did not understand exactly what to do, the test supervisor explained the test again and gave an example. This was done in accordance with the test protocols described in Platvoet et al. (2018) and Hoeboer et al. (2018). After completing a test, children were asked to go to another test that they had not yet performed.

TABLE 1 | Raw scores (median with minimum and maximum value) of children from grades 3 through 8 on the four subtests of the KTK3+ (number of successful attempts) and the AST (time in seconds).

		N	Age	KTK3+				AST
				Balance	Jumping sideways	Moving sideways	Eye-hand coordination	
Grade 3	Total	467	6 [5–8]	24 [2–72]	36 [16–69]	31 [15–46]	2 [0–35]	39.3 [22.2–59.7]
	Girls	254	6 [5–8]	26 [5–72]	35 [18–69]	31 [18–44]	1 [0–29]	40.7 [25.1–59.7]
	Boys	213	6 [5–7]	21 [2–62]	38 [16–63]	31 [15–46]	4 [0–35]	37.4 [22.2–58.4]
Grade 4	Total	517	7 [6–9]	29 [5–65]	47 [12–83]	36 [14–54]	14 [0–58]	32.9 [18.1–59.9]
	Girls	245	7 [6–9]	33 [6–65]	47 [12–79]	36 [19–52]	11 [0–45]	34.2 [22.0–58.1]
	Boys	272	7 [6–9]	27 [5–62]	47 [18–83]	36 [14–54]	16 [0–58]	31.5 [18.1–59.9]
Grade 5	Total	520	8 [7–10]	33 [0–68]	54 [12–85]	39 [15–58]	8 [0–51]	29.9 [18.4–57.7]
	Girls	272	8 [7–10]	36 [7–68]	53 [22–83]	39 [18–58]	5 [0–51]	30.5 [19.5–57.7]
	Boys	248	8 [7–10]	31 [0–67]	55 [12–85]	39 [15–57]	13 [0–42]	29.5 [18.4–54.1]
Grade 6	Total	515	9 [8–12]	39 [8–72]	60 [25–99]	43 [18–65]	16 [0–52]	30.2 [18.0–56.2]
	Girls	264	9 [8–10]	43 [9–72]	60 [25–99]	43 [18–63]	12 [0–52]	31.0 [18.6–56.2]
	Boys	251	9 [8–12]	35 [8–68]	61 [31–96]	43 [18–65]	23 [0–51]	29.1 [18.0–54.4]
Grade 7	Total	543	10 [9–12]	42 [11–72]	65 [26–109]	45 [19–66]	24 [0–74]	30.0 [16.2–59.3]
	Girls	295	10 [9–12]	45 [12–72]	64 [26–109]	46 [23–65]	18 [0–59]	31.3 [16.2–57.6]
	Boys	248	10 [9–12]	38 [11–66]	67 [36–103]	45 [19–66]	31 [0–74]	28.0 [16.7–59.3]
Grade 8	Total	545	11 [10–13]	44 [9–72]	71 [31–106]	47 [21–75]	32 [0–69]	26.8 [17.5–51.7]
	Girls	295	11 [10–13]	47 [16–72]	70 [38–106]	47 [24–75]	29 [0–67]	27.8 [17.7–51.7]
	Boys	250	11[10–13]	43 [9–72]	73 [31–101]	47 [21–71]	36 [1–69]	25.5 [17.5–50.7]

TABLE 2 | Pearson correlation coefficients between the KTK3+ and the AST.

		KTK3+						
		Overall	Grade 3	Grade 4	Grade 5	Grade 6	Grade 7	Grade 8
AST	Overall	0.512						
	Grade 3		0.467					
	Grade 4			0.485				
	Grade 5				0.508			
	Grade 6					0.597		
	Grade 7						0.520	
	Grade 8							0.490

Children who completed all tests within the 45-min PE lesson were given the opportunity to play freely in a separate part of the gym under the supervision of their PE teacher.

Data Analysis

First, descriptives of raw scores for boys and girls were calculated for each test and presented for each grade. Next, raw test scores for each individual were transformed into percentile scores corrected for gender and grade (instead of the classical chronological age comparisons, e.g., grade 3 boys, grade 3 girls, grade 4 boys). This resulted in five percentile scores for each child (i.e., four for the KTK3+ and one for the AST). Subsequently, the four percentile scores for the KTK3+ were summed to calculate an overall percentile score for the KTK3+. This resulted in two percentile scores per child: first, a percentile score based on the KTK3+ and second, a percentile score based on the AST. Bivariate Pearson correlation analysis was used to determine the association between both scores. Additionally, the same analysis (from raw scores to one percentile score) was conducted to determine the correlation between the KTK3 (without hand-eye subtest) and AST, since the

AST does not consist an object control skill. A correlation coefficient <0.30 can be interpreted as negligible, 0.30–0.50 as low, 0.51–0.70 as moderate, 0.71–0.90 as high and >0.90 as very high (Hinkle et al., 2003).

For the KTK3+ and the AST, children were classified into five categories based on the calculated percentile scores: <5, 5–15, 16–85, 86–95 or >95%. To assess the agreement between the KTK3+ and AST, cross tabs between both tests were used and Pearson Chi square (Chi^2) and weighted Cohen's Kappa (k) values were calculated. Using the weighted Cohen's Kappa takes the size of differences in classification into account (e.g., differences between classification <5 and 5–15% have less impact on Cohen's Kappa than between <5 and >95%). The following cut-off scores were used: A Kappa of 0–0.20 was interpreted as low, 0.21–0.40 as fair, 0.41–0.60 as moderate, 0.61–0.80 as substantial and >0.80 as an almost perfect agreement (Landis and Koch, 1977). The significance level was set at $p < 0.05$ for all analyses. The statistical analysis was first performed on all participants, and thereafter repeated for each individual grade.

TABLE 3 | Number of children in each motor skills proficiency level per test Chi^2(25) = 897.5, p <0 .001.

		KTK3+					Total
		<5%	5–15%	15–85%	85–95%	95–100%	
AST	<5%	33	43	70	4	1	151
	5–15%	39	66	197	7	0	309
	15–85%	77	195	1,623	196	85	2,176
	85–95%	0	7	199	65	35	306
	95–100%	0	0	86	39	40	165
Total		149	311	2,175	311	161	3,107

TABLE 4 | Agreement scores and Cohen's Kappa between the KTK3+ and the AST.

	Agreement (%)	Chi^2	Cohen's Kappa	p-value
Overall	58.8	583.8	0.150	<0.001
Grade 3	56.5	64.2	0.103	<0.001
Grade 4	58.4	83.7	0.138	<0.001
Grade 5	60.2	132.4	0.183	<0.001
Grade 6	60.4	134.5	0.182	<0.001
Grade 7	59.1	152.3	0.159	<0.001
Grade 8	58.0	97.3	0.132	<0.001

RESULTS

KTK3+ and AST Scores by Grade and Gender

Mainly due to injury and absence during measurements, as well as some technical and measurement errors (i.e., unsaved data due to a bad internet connection), 3,107 children (of which 1,641 are girls) of the total sample of 3,803 were included in the agreement analysis. The descriptive statistics of the raw scores on the four subtests of the KTK3+ and AST are shown in **Table 1**.

Correlation Between the KTK3+ and AST

A positive and significant correlation coefficient was found between the KTK3+ and AST (both corrected for grade and gender; r = 0.51, p < 0.001). Additionally, the correlation coefficient between the KTK3 (without the hand-eye coordination test) and the AST was positive and significant (r = 0.53; p < 0.001). Within each grade, the strongest correlation coefficient between KTK3+ and AST was found in grade 6 (r = 0.60, p < 0.001), while the weakest, but still significant, correlation coefficients was found in grade 3 (0.47, p < 0.001, see **Table 2**).

Agreement Between the KTK3+ and AST

The cross tab between the KTK3+ and AST showed a fair level of agreement (Chi^2 = 583,7, p < 0.001, weighted k = 0.26, agreement of 58.8%). The cross tabs between the KTK3+ and the AST within each grade showed low to moderate levels of agreement, from agreement of 56.5% in grade 3 (Chi^2=64.2; weighted k = 0.19) to agreement of 60.4% in grade 6 (Chi^2 = 134.5; weighted k = 0.30). The number of children classified in each percentile category for each test can be found in **Table 3**, whereas the agreement percentages and weighted k for each grade are presented in **Table 4**. Nineteen children (0.6%) were classified in one of the two lowest categories on the first test while they were classified in one of the two highest categories on the other test.

DISCUSSION

In recent years, the KTK3+ and the AST were developed to assess children's FMS in applied settings. This study found a moderate correlation coefficient (r = 0.51) between those tests,

ranging from r = 0.47 (grade 3) to r = 0.60(grade 6). The association is in line with previous studies that compared motor skill tests (Smits-Engelsman et al., 1998; Cools et al., 2010; Fransen et al., 2014). Additionally, the main aim of this study was to determine the agreement in classifying children's FMS level with the KTK3+ and the AST in 6- to 12-year-olds. The classification analysis showed a fair agreement between the tests (level of agreement: 58.8% with k = 0.26). Also, major discrepancies in classification between the tests were found in only 19 (0.6%) of the children. So, although we only found a fair agreement, the chance that a child was 'misclassified' with one of the two tests is small.

The AST was originally validated by studying its relationship with the original KTK from Kiphard and Schilling (1974), Kiphard and Schilling (2007). The validity and reliabiltiy of the original and modified KTK have been shown before (Kiphardt and Schilling, 1974; Kiphardt and Schilling, 2007; Platvoet et al., 2018; Vandorpe et al., 2012; D'Hondt et al., 2013). Hoeboer and colleagues (2016) found overall correlations for AST-1 and AST-2 of r = −0.47 and r = −0.50, respectively. In their second validation study, in which the test was modified for different age groups, the authors found a higher correlation between the AST and KTK (i.e., AST-2: r = −0.65, AST-3: r = −0.60) (Hoeboer et al., 2018). The correlation (r = 0.51) found in this study, which used the age-specific AST-2 and AST-3, is lower than that found by Hoeboer and colleagues (2018). The inclusion of an object control test in the KTK3+ may explain the lower correlation, as object control is not included in the AST-2 and AST-3. The strength of the association in this study is indeed comparable to the data in the study of Hoeboer and colleagues in 2016, which used tracks that included object control. However, without the object control test item in the KTK3+, the correlation coefficient (r = 0.53) between the age-specific AST and KTK3+ does not enhance the strength of the association found in the study by Hoeboer and colleagues (2018). The inclusion of object control in the KTK3+ was not the only difference between the tests used in this study. In the KTK3+ the test item 'hopping' was excluded for several reasons (mainly feasibility and risk of injury; Novak et al., 2017), which may explain differences in the association found. Nevertheless, the association found in this study is in line with results from other studies that compared different motor skill tests (Smits-Engelsman et al., 1998; Cools et al., 2010; Fransen et al., 2014). However, in for example the study of Fransen et al. (2014) the KTK was compared with the BOT-2 short form which

also measures fine motor skills. That explanation cannot be used in this study as both the KTK3+ and the AST aim to measure gross motor skills.

Although associations between tests are often used to compare and validate tests, this methodology is less suitable for determining whether children are classified identically based on different tests. This is important, especially in PE, as these kinds of tests are often used to classify children as having a motor disorder or an excellent motor competence. Therefore, the agreement between the tests was considered. The overall level of agreement between both tests is fair (level of agreement: 58.8% with $k = 0.26$). The agreement between the KTK3+ and the AST was lower than previously found in a study comparing the KTK and the Bruininks–Oseretsky Test of Motor Proficiency second edition (BOT-2; Fransen et al., 2014) and in a study in which 300 Portuguese children were compared with the Movement-Assessment Battery for Children and the KTK (Mendes, et al., 2009). The authors found an agreement of 78% with a Cohen's kappa of 0.65. An explanation for the higher agreement between both tests (BOT-2 and M-ABC) with the KTK than between the KTK3+ and AST may be the use of only one outcome measures, which might hamper the construct validity (Scheuer et al., 2019). For the KTK3+ and AST, one outcome measure is presented, though the one outcome measure for the KTK3+ is a composite score of four separate tests whereas one score is used for the AST. A fair agreement may also suggest that the convergent validity of both tests is fair to low. Although both tests reflect children's proficiency in FMS, different skills are used in both tests. This suggests that measuring different fundamental movement skills, may not reflect all or similar aspects of FMS, nor motor competence (Cools et al., 2010; Hulteen, et al., 2020).

The skills and properties of each test may also impact the agreement and association between the tests. As mentioned before, the KTK3+ includes skills of all three categories of the FMS, whereas the AST does not include an object control skill. Furthermore, personal characteristics of the children may impact the execution of the tests. It might be that some children, due to their personality, execute a test exactly as instructed (i.e., as well as possible instead of as quickly as possible) while others may perform the test as quickly as possible. Individual performances might therefore be influenced by a speed-accuracy trade-off (Schmidt et al., 1985), which influences those performances in which speed plays an important role (i.e., KTK jumping sideways, KTK moving sideways, and the AST). This effect may be more profound in the AST than in the KTK3+ because the KTK3+ also include subtests with less time pressure (especially balancing). Additionally, physical fitness may be part of the test outcomes to different degrees. The KTK3+ is known to have only limited interference from physical fitness (Vandorpe et al., 2011). In this study, boys outperform girls on the tracks and boys, in general, have better physical fitness (Roth et al., 2018; Lisowski et al., 2020). This might suggest that physical fitness plays a more important role in the outcome of the AST.

The results of this study support the complexity of the construct of motor competence, and the difficulty to measure

the construct with one easy to use test. This is in line with recent developments in literature (Logan et al., 2018; Hulteen et al., 2020). The results at least suggest that both tests seem to measure different parts of FMS i.e., different skills. Therefore, we should realize that test results provide an indication of children's proficiency related to the specific tasks performed (Van Waelvelde et al., 2007), but they may not necessarily provide an accurate indication for all FMS, let alone motor competence. Thus, the KTK3+ provides an indication of how children score on balancing backward, jumping sideways, moving sideways, and throwing and catching a tennis ball, while the AST provides information about how fast children can perform a diverse range of movements (e.g., traveling jumps, alligator crawl (backward), bunny hopping, and forward rolling) over a track. A probably more valuable factor for PE is the ecological validity of both tests, i.e., the extent to which both tests are predictors for children's physical activity levels and sport performance in later stages of life. This should be taken into account into further research.

Although several explanations can be given for the differences between the tests, it is also important to note that only seven children with a low overall score (<15%) on the KTK3+ had a high score on the AST (>85%) and 12 children with a high score on the KTK3+ scored low on the AST, representing 0.6% of all children. Although the agreement between the tests is fair, the chance that a child was 'misclassified' completely is small. However, PE teachers should always use their own observations too and realize that the results of valid and reliable tests also have their limitations, which can lead to misinterpretations of children's actual proficiency.

This study has strengths and some limitations too. The main strengths of this study are its large sample size and the broad range of ages. Most comparable studies (e.g., Cools et al., 2010; Zuvela et al., 2011; Lane and Brown, 2015) used sample sizes between 30 and 100 participants. This study assessed more than 3,000 children from 21 schools (with a professional PE teacher), which resulted in a representative sample for the Dutch primary school population from grades 3 through 8. A limitation of this study is that we do not have information about children's sport participation and physical fitness. This information could probably have helped us to explain the differences in classification between both tests. Therefore, more research is warranted. Second, the KTK3+ and the AST are developed to measure children's FMS. In literature there is still discussion what FMS (e.g., Barnett et al., 2016) are and how FMS relate to the broader construct of motor competence. Also, it suggests that, although the KTK3+ and AST have been validated, research remains needed to determine especially the construct validity (i.e., content, internal structure, convergent, and discriminant validity) of both tests.

To conclude, we found a moderate correlation between the KTK3+ and AST, indicating that both tests seem to measure a similar construct to some extent. However, we found a fair agreement when classifying children into five categories based on their test scores. This indicates that the choice of test determines how a child is classified, although the percentage of children completely misclassified is small. Practitioners should carefully think about the purpose of an FMS assessment (e.g.,

screening and monitoring individuals or groups of children or evaluating interventions or benchmarking). Based on their considerations, and in combination with practical issues such as costs, available materials and time to execute, they can choose the most appropriate assessment tool (Cools et al., 2009; Scheuer et al., 2019). Moreover, practitioners should realize that the outcome of a test provides them with an insight into how well a child performs the activities in that test. Furthermore, researchers should conduct longitudinal studies to determine the predictive value of current test results for future physical activity levels and level of sport performance.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the HAN University of Applied Sciences Ethical Committee. Written informed consent for participation was not provided by the participants' legal guardians/next of kin because the Ethical Committee approved a passive informed consent, which means all legal guardians were informed about the study. If a legal guardian did not want to participate in the study, data were not collected of the specific children.

AUTHOR CONTRIBUTIONS

All authors contributed to conception and design of the study. MN, SP, JH, and AW organized data collection. MN organized the database. MN and SP performed the statistical analysis. MN and SP wrote the first draft of the manuscript. All authors contributed to manuscript revision, read, and approved the submitted version.

FUNDING

ACKNOWLEDGMENTS

REFERENCES

Barnett, A., and Peters, J. (2004). "Motor proficiency assessment batteries," in *Developmental motor disorders: a neuropsychological Perspective*. Editors D. Dewey and D. E. Tupper (New York, NY: Guilford), 67–109.

Barnett, L. M., Stodden, D., Cohen, K. E., Smith, J. J., Lubans, D. R., Lenoir, M., et al. (2016). Fundamental movement skills: an important focus. *J. Teach. Phys. Educ.* 35, 219–225. doi:10.1123/jtpe.2014-0209

Bruininks, R. H., and Bruininks, B. D. (2005). *Bruininks–Oseretsky test of motor proficiency, (BOT-2)*. Minneapolis, MN: Pearson Assessment.

Cattuzzo, M. T., dos Santos Henrique, R., Ré, A. H., de Oliveira, I. S., Melo, B. M., de Sousa Moura, M., et al. (2016). Motor competence and health related physical fitness in youth: a systematic review. *J. Sci. Med. Sport.* 19, 123–129. doi:10.1016/j.jsams.2014.12.004

Clark, J. E. (2007). On the problem of motor skill development. *J. Phys. Educ. Recreat. Dance.* 78, 39–44. doi:10.1080/07303084.2007.10598023

Cliff, D. P., Okely, A. D., Smith, L. M., and McKeen, K. (2009). Relationships between fundamental movement skills and objectively measured physical activity in preschool children. *Pediatr. Exerc. Sci.* 21 (4), 436–449. doi:10.1123/pes.21.4.436

Cools, W., De Martelaer, K., Vandaele, B., Samaey, C., and Andries, C. (2010). Assessment of movement skill performance in preschool children: convergent validity between MOT 4-6 and M-ABC. *J. Sports Sci. Med.* 9 (4), 597–604.

Cools, W., Martelaer, K., Samaey, C., and Andries, C. (2009). Movement skill assessment of typically developing preschool children: a review of seven movement skill assessment tools. *J. Sports Sci. Med.* 8, 154–168.

D'Hondt, E., Deforche, B., Gentier, I., De Bourdeaudhuij, I., Vaeyens, R., Philippaerts, R., et al. (2013). A longitudinal analysis of gross motor coordination in overweight and obese children versus normal-weight peers. *Int. J. Obes.* 37 (1), 61–67. doi:10.1038/ijo.2012.55

Faber, I. R., Oosterveld, F. G., and Nijhuis-Van der Sanden, M. W. (2014). Does an eye-hand coordination test have added value as part of talent identification in table tennis? A validity and reproducibility study. *PLoS One.* 9, e85657. doi:10.1371/journal.pone.0085657

Fransen, J., D'Hondt, E., Bourgois, J., Vaeyens, R., Philippaerts, R. M., and Lenoir, M. (2014). Motor competence assessment in children: convergent and discriminant validity between the BOT-2 Short Form and KTK testing batteries. *Res. Dev. Disabil.* 35, 1375–1383. doi:10.1016/j.ridd.2014.03.011

Gallahue, D. L., Ozmun, J. C., and Goodway, J. D. (2012). Understanding motor development: Infants, children, adolescents, adults (7th Edn.). New York, NY: McGraw-Hill.

Gallahue, D. L., and Donnelly, F. C. (2003). Developmental physical education for all children. Champaign, IL: Human Kinetics.

Henderson, S. E., Sugden, D. A., and Barnett, A. L. (2008). *M-ABC-2: movement assessment battery for Children-2*. London, UK: Harcourt.

Herrmann, C., and Seelig, H. (2017). Basic motor competencies of fifth graders. *Ger. J. Exerc. Sport Res.* 47 (2), 110–121. doi:10.1007/s12662-016-0430-3

Hinkle, D. E., Wiersma, W., and Jurs, S. G. (2003). *Applied statistics for the behavioral sciences*. Boston, MA: Houghton Mifflin.

Hoeboer, J., De Vries, S., Krijger-Hombergen, M., Wormhoudt, R., Drent, A., Krabben, K., et al. (2016). Validity of an athletic skills track among 6- to 12-year-old children. *J. Sports Sci.* 34 (21), 2095–2105. doi:10.1080/02640414.2016.1151920

Hoeboer, J., Krijger-Hombergen, M., Savelsbergh, G., and De Vries, S. (2018). Reliability and concurrent validity of a motor skill competence test among 4- to 12-year old children. *J. Sports Sci.* 36 (14), 1607–1613. doi:10.1080/02640414.2017.1406296

Hulteen, R. M., Barnett, L. M., True, L., Lander, N. J., del Pozo Cruz, B., and Lonsdale, C. (2020). Validity and reliability evidence for motor competence assessments in children and adolescents: a systematic review. *J. Sports Sci.* 38, 1717–1798. doi:10.1080/02640414.2020.1756674

Kiphard, E. J., and Schilling, F. (1974). *KörperkoordinationsTest für Kinder.* Weinheim: Beltz Test GmbH.

Kiphard, E. J., and Schilling, F. (2007). *Körperkoordinationstest für Kinder: KTK.* Weinheim: Beltz Test GmbH.

Klingberg, B., Schranz, N., Barnett, L. M., Booth, V., and Ferrar, K. (2018). The feasibility of fundamental movement skill assessments for pre-school aged children. *J. Sports Sci.* 37 (4), 378–386. doi:10.1080/02640414.2018.1504603

Landis, J. R., and Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics.* 33, 159–174. doi:10.2307/2529310

Lane, H., and Brown, T. (2015). Convergent validity of two motor skill tests used to assess school-age children. *Scand. J. Occup. Ther.* 22 (3), 161–172. doi:10.3109/11038128.2014.969308

Lisowski, P., Kantanista, A., and Bronikowski, M. (2020). Are there any differences between first grade boys and girls in physical fitness, physical activity, BMI, and sedentary behavior? Results of HCSC study. *Ijerph.* 17, 1109. doi:10.3390/ijerph17031109

Lloyd, M., Saunders, T. J., Bremer, E., and Tremblay, M. S. (2014). Long-term importance of fundamental motor skills: a 20-year follow-up study. *Adapt. Phys. Activ. Q.* 31, 67–78. doi:10.1123/apaq.2013-0048

Logan, S. W., Ross, S. M., Chee, K., Stodden, D. F., and Robinson, L. E. (2018). Fundamental motor skills: a systematic review of terminology. *J. Sports Sci.* 36 (7), 781–796. doi:10.1080/02640414.2017.1340660

Longmuir, P. E., Boyer, C., Lloyd, M., Borghese, M. M., Knight, E., Saunders, T. J., et al. (2017). Canadian Agility and Movement Skill Assessment (CAMSA): validity, objectivity, and reliability evidence for children 8-12 years of age. *J. Sport Health Sci.* 6 (2), 231–240. doi:10.1016/j.jshs.2015.11.004

Loprinzi, P. D., Davis, R. E., and Fu, Y. C. (2015). Early motor skill competence as a mediator of child and adult physical activity. *Prev. Med. Rep.* 2, 833–838. doi:10.1016/j.pmedr.2015.09.015

Lubans, D. R., Morgan, P. J., Cliff, D. P., Barnett, L. M., and Okely, A. D. (2010). Fundamental movement skills in children and adolescents: review of associated health benefits. *Sports Med.* 40, 1019. doi:10.2165/11536850-000000000-00000

Mendes, R. S., Fonseca, P. M., and Rodrigues, L. P. (2009). Relationship between two motor coordination tests: the movement assessment battery for children (M-ABC) and the Körperkoordinationstest für Kinder (KTK). Proceedings of the NASPSPA conference (Austin, Texas). *J. Sport Exerc. Psy.* 31, S38.

Morley, D., Van Rossum, T., Richardson, D., and Foweather, L. (2019). Expert recommendations for the design of a children's movement competence assessment tool for use by primary school teachers. *Eur. Phys. Educ. Rev.* 25 (2), 524–543. doi:10.1177/1356336X17751358

Novak, A. R., Bennett, K. J. M., Beavan, A., Pion, J., Spiteri, T., Fransen, J., et al. (2017). The applicability of a short form of the körperkoordinationstest für kinder for measuring motor competence in children aged 6 to 11 years. *J. Mot. Learn. Dev.* 5, 227–239. doi:10.1123/jmld.2016-0028

Platvoet, S., Faber, I. R., de Niet, M., Kannekens, R., Pion, J., Elferink-Gemser, M. T., et al. (2018). Development of a tool to assess fundamental movement skills in applied settings. *Front. Educ.* 3, 75. doi:10.3389/feduc.2018.00075

Roth, A., Schmidt, S. C. E., Seidel, I., Woll, A., and Bös, K. (2018). Tracking of physical fitness of primary school children in trier: a 4-year longitudinal study. *Biomed. Res. Int.* 2018, 7231818. doi:10.1155/2018/7231818

Scheuer, C., Herrmann, C., and Bund, A. (2019). Motor tests for primary school aged children: a systematic review. *J. Sports Sci.* 37 (10), 1097–1112. doi:10.1080/02640414.2018.1544535

Schmidt, R. A., Sherwood, D. E., Zelaznik, H. N., and Leikind, B. J. (1985). "Speed-accuracy trade-offs in motor behavior: theories of impulse variability," in *Motor behavior.* Editors H. Heuer, U. Kleinbeck, and K. H. Schmidt (Berlin: Heidelberg), 79–123.

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. *Hum. Mov. Sci.* 17, 699e709. doi:10.1016/s0167-9457(98)00019-0

Stodden, D. F., Goodway, J. D., Langendorfer, S. J., Roberton, M. A., Rudisill, M. E., Garcia, C., et al. (2008). A developmental perspective on the role of motor skill competence in physical activity: an emergent relationship. *Quest.* 60, 290–306. doi:10.1080/00336297.2008.10483582

Tieman, B. L., Palisano, R. J., and Sutlive, A. C. (2005). Assessment of motor development and function in preschool children. *Ment. Retard. Dev. Disabil. Res. Rev.* 11 (3), 189–196. doi:10.1002/mrdd.20074

Ulrich, D. A. (1985). *Test of gross motor development.* Austin, TX: Pro-ED. Inc.

Ulrich, D. A. (2000). "Test of gross motor development" in *Examiner's manual.* 2nd Edn. (Austin, TX: Pro-ED. Inc).

Van Waelvelde, H., Peersman, W., Lenoir, M., and Engelsman, B. C. M. S. (2007). Convergent validity between two motor tests: movement-ABC and PDMS-2. *Adapt. Phys. Act. Quart.* 24 (1), 59–69. doi:10.1123/apaq.24.1.59

Vandorpe, B., Vandendriessche, J., Lefèvre, J., Pion, J., Vaeyens, R., Matthys, S., et al. (2011). The KörperkoordinationsTest für Kinder: reference values and suitability for 6-12-year-old children in Flanders. *Scand. J. Med. Sci. Sports.* 21, 378–388. doi:10.1111/j.1600-0838.2009.01067.x

Vandorpe, B., Vandendriessche, J., Vaeyens, R., Pion, J., Matthys, S., Lefevre, J., et al. (2012). Relationship between sports participation and the level of motor coordination in childhood: a longitudinal approach. *J. Sci. Med. Sport.* 15 (3), 220–225. doi:10.1016/j.jsams.2011.09.006

Wiart, L., and Darrah, J. (2001). Review of four tests of gross motor development. *Dev. Med. Child. Neurol.* 43, 279–285. doi:10.1017/s0012162201000536

Yoon, D., Scott, K., Hill, M., Levitt, N. S., and Lambert, E. V. 2006). Review of three tests of motor proficiency in children. *Percept. Mot. Skills.* 102, 543–551. doi:10.2466/pms.102.2.543-551

Zuvela, F., Bozanic, A., and Miletic, D. (2011). POLYGON—a New fundamental movement skills test for 8 Year old children: construction and validation. *J. Sports Sci. Med.* 10 (1), 157–163.