The convergent validity of four fine motor assessment tools in 5 to 12 years old children

Simons, Johan¹, Van Damme, Tinne² & Franssen Willem³

¹Faculty of Kinesiology and Rehabilitation Sciences, Department of Rehabilitation Sciences, KU Leuven, O & N IV, Bus 1510, B 3000 Leuven, Belgium, Johan.simons@kuleuven.be
²Faculty of Kinesiology and Rehabilitation Sciences, Department of Rehabilitation Sciences, KU Leuven, O & N IV, Bus 1510, B 3000 Leuven, Belgium University Psychiatric Centre, Campus Leuven, Gasthuisberg, Child & Adolescent Psychiatry, Herestraat 49, B 3000 Leuven, Belgium, Tine.vandamme@kuleuven.be
³Faculty of Kinesiology and Rehabilitation Sciences, Department of Rehabilitation Sciences, KU Leuven, O & N IV, Bus 1510, B 3000 Leuven, Belgium

Abstract

The aim of this longitudinal study was to evaluate the convergent validity of four fine motor assessment tools in 5 to 12 year old children. Other research goals were the investigation of age and gender differences concerning fine motor skills in this population. In this longitudinal study, children from the last year of nursery school were followed till their sixth grade. During these 7 years, four fine motor assessment tools were administered: The Movement-ABC, the Punktiertest für Kinder, the Epreuve de pointillage and the Gibson Spiral Maze. Generally, low correlations were found between the four assessment tools. The correlation between the Epreuve de pointillage and the Punktiertest für Kinder was the highest. The lowest correlations were found with the Movement-ABC. Secondly, significant sex differences were observed. Improvements in fine motor skills with age were noted, especially in the lower grades. Finally, girls performed significantly better than boys on the fine motor assessment tools. The results of the Epreuve de pointillage and the Punktiertest für Kinder indicate that they both measure a similar aspect of fine motor skills but different aspects were measured in the other instruments. Overall, more research on this topic is needed.

Keywords: Movement-ABC, Epreuve de pointillage, Gibson Spiral Maze, Punktiertest für Kinder, fine motor skills, convergent validity, fine motor assessment tools, children, gender, age, longitudinal study

Introduction

In everyday life, a person makes countless fine motor actions: grasping for a pen, texting, eating, etc. To fulfil these fine motor actions, a person needs Fine Motor Skills (FMS). Children fill 46% of their time in nursery school with fine motor (FM) activities.
(Marr, Cermak, Cohn & Henderson, 2003) and adults execute several of these FM activities while performing their job e.g. dentist (Wang et al., 2011). Burton and Miller (1998) defined these FMS as “the ability to perform movements with the smaller muscles, usually involving precise manipulation of small objects with the hands and fingers” (p. 365). In the literature, the definitions of FMS and manual dexterity are frequently interchangeable. Wang et al. (2011) defined manual dexterity as “the ability to coordinate and manipulate objects in a timely manner” (p. 313). Bernstein (1996) defined dexterity not as a skill or as a combination of skills, but as the ability to find a motor solution for any external situation quickly, economically, correctly, expeditiously and resourcefully. In this study, the definition of Bernstein (1996) was used to define manual dexterity.

In the past few decades several taxonomies tried to classify the abilities that are involved in psychomotor activities, which Burton and Miller (1998) termed as motor abilities. These taxonomies go beyond the distinction between fine and gross motor skills. The taxonomy of Fleishman and Ellison (1962) divided the abilities in two main categories: perceptual motor abilities and physical proficiency abilities. Only the classification of the perceptual motor abilities was used for this study because the physical proficiency abilities are irrelevant in FMS. The eleven perceptual motor abilities for FMS are: control precision, multi-limb coordination, response orientation, reaction time, speed of limb movement, rate control, manual dexterity, finger dexterity, arm-hand steadiness, wrist finger speed and aiming (Fleishman & Ellison, 1962). Burton and Miller (1998) published their taxonomy of the concept of general motor ability in 1998. In this taxonomy, there is a theoretical hierarchy in which three different concepts are placed. At the basic level, there is the “general motor ability”, a single trait of an individual that determines the performance of all movement skills (Schmidt & Lee, 1999). The next level is the “motor abilities”, defined by Burton and Miller (1998) as “general traits or capacities of an individual that underlie the performance of a variety of movement skills” (p. 43). At the highest level of the taxonomy of Burton and Miller (1998) there are the “movement skills”, defined as “a specific class of goal-
directed movement patterns such as throwing, running, writing, speaking, driving and hammering” (p. 44). Burton and Rodgerson (2001) build up a new taxonomy of movement skills and general motor ability because of the inconsistent use of terms and constructs, and a low within-task correlation between different assessment tools. Their taxonomy consists of four levels: movement skills, movement skill sets, movement skill foundations and general movement ability (Burton & Rodgerson, 2001). The movement skill sets in this new taxonomy are the movement skills that are clustered together into a subtest in an assessment tool. Simons (2014) prefers a more practical classification of the motor skills after the first years of life. He distinguishes four categories that can be made: basic motor skills, perceptual-motor skills, physical fitness and specific skills. In this classification the FMS can be included in the perceptual motor skills category.

Injuries, diseases, developmental problems, age, gender and motivation of the individual are parameters that influence the FMS (Backman, Gibson, Cork & Parsons, 1992). In typical developing children, the level of movement skill performance improves with age (Espenschade & Eckert, 1980; Rarick, 1981). Just like the basic movement skills, the perceptual motor skills will develop in children from 2 to 7 years old (Gallahue & Ozmun, 2006; Sherril, 2004). The study of Humprey, Jewell and Rosenberger (1995) with 184 children between 2 and 7 years old, affirmed that there is a positive linear relationship between the FMS and the maturation in children. The studies of Pehoski, Henderson and Tickle-Degnen (1997a, 1997b) also came to this conclusion. The study of Smith, Hong and Presson (2000) found a linear increase of FM dexterity in children across ages 5 to 10 and found a curvilinear relationship in children across ages 10 to 20 years.

Besides age, gender could be an important parameter that influences FMS. Malina and Bouchard (1991) concluded that boys were better in jumping skills, running and throwing skills, and girls were better in FM abilities e.g. manual dexterity. The opposite of this conclusion was found in the study of Ikeda and Aoyagi (2009). Other studies (e.g., Bruininks, 1978; Stein & Yerxa,
1990; Pehoski et al., 1997a, 1997b; O’Neill, 1995; Folio & Fewell, 1983) concluded that there was no statistical gender difference in manual dexterity between boys and girls (Smith et al., 2000). Because most motor assessment tools are not equipped with separated norm tables according to gender, gender differences can be an important area of concern (Park & Simons, 2012).

In elementary aged school children, major difficulties in FMS can be identified in 8-15% of the population (Cratty, 1986). Consequently, the assessment of FMS in school-aged children is a very important task (Backman et al., 1992). The abundance of FMS assessment tools causes a lack of consensus (Cools, Martelaer, Samaey & Andries, 2009) on how FMS across all ages must be measured. This abundance is shown in the reviews of Yancosek and Howell (2009), Schoneveld, Wittink and Takken (2009) and Cools et al. (2009), in which there are already 18 different FMS assessment tools found. The selection of an appropriate assessment tool for a specific population (e.g. elementary school-aged children) is a very important issue because of the abundance. In the review of Tieman, Palisano and Sutlive (2005) five questions were described to facilitate this selection procedure: who, what, why (purpose: evaluative…) and where we examine. The last question asks for the external constraints of the test. These constraints are: the financial costs, time costs, practicality and familiarity, etc. (Yancosek & Howell, 2009).

There is a need for more comparison studies among the assessment tools. To our knowledge, there are thirteen comparison studies of several FMS assessment tools. Five of these were correlation studies between assessments with a fine and a gross motor subscale. The comparisons that were made weren’t solely of the FMS subscale. There are only two studies (Van Waelvelde, Peersman, Lenoir & Smits-Engelsman, 2007; Van Hartingsveldt, Cup & Oostendorp, 2005) where there is a comparison of only the FMS subscale of these global motor assessments. In these two studies they compared the Movement-ABC (M-ABC) that measures the perceptual-motor skills with the Peabody Developmental Motor Scales-2 (PDMS-2) that measures the motor...
development of the children. The FM construct of the PDMS-2 consists of grasping and visual motor integration activities, while the FM construct of the M-ABC consists of items concerning manual dexterity. The correlations that were found (r=0.48 and r=0.69) indicated a moderate agreement between the tests. Differences in the sample and/or sampling error could explain the differences. Beside the comparison of these global assessments with a FM subscale, there is need for comparison studies of the assessments that only measure FMS or an aspect of it.

The aim of this longitudinal study was to evaluate the convergent validity of four fine motor assessment tools in 5 to 12-year-old children. This relationship shows whether the different tests measure the same aspects of FMS. It is expected that the four instruments measured the same aspect. Secondly, the relationship between FMS and age was investigated. It was expected that the results would confirm the assumption of Smith et al. (2000), Humphrey et al., (1995) and Pehoski et al. (1997a, 1997b), that FMS will increase when children get older. Thirdly, the difference in FMS depending on gender will be evaluated. It was expected that there would be no statistical difference between the results of boys and girls (Bruininks, 1978; Stein & Yerxa, 1990; Pehoski et al., 1997a, 1997b; O’Neill, 1995; Folio & Fewell, 1983).

Method

Participants

In a school for typical developing children, children were followed from the last year of nursery school (5 years old) till their last year of elementary school (12 years old). During these 7 years, a test battery of several assessments was conducted annually, in the same time of the year. Children’s socioeconomic status (SES) was determined by using postcode as a proxy measure accordingly to census data. Most children came from a family with higher socio-economic class. The parents/caregivers were informed about the aim of the project and how the confidentiality of
their personal information would be handled. Before the children were included, the parents and the school board had to give written informed consent as the children were not old enough to provide informed consent. Exclusion criteria were: a reported neurological, motor problem and/or learning disability. In each subsequent grade, the children were included in this study when the data of all tests were available.

Only 22 children (10 boys and 12 girls) of the 75 who participated over the years, had a complete data set from preschool until sixth grade. Their mean age was 8.52 years with a standard deviation of 2.04 years old. The minimum age was 4.10 years and the maximum was 12.1-year-old. A t-test between the age of the experimental group (the group who completed all the test) and the others was non-significant \( (t= 0.375, p=0.708) \), of the 22 children, 18 were right handed and only 4 were left handed.

**Materials**

To investigate the FMS in these children, four FM assessment tools (for children), which are most commonly used in Flanders, were applied: The Epreuve de pointillage (Stambak, 1966), the Punktiertest für Kinder (PTK) (Schilling, 2009), the Gibson Spiral Maze (GSM) (Gibson, 1964) and the Movement-Assessment Battery for Children (M-ABC) (Henderson & Sugden, 1992).

The Epreuve de pointillage is a test for children from 6 to 14 years old, designed by Stambak in 1966. This is a test, measuring speed combined with precision. During the test, the children have to draw a line in as many squares \( (1\text{cm}^2) \) as possible in 1 minute. This French test is normed on 385 French children (6-14 years old). The reliability of the Epreuve de pointillage is \( \text{rtt}=0.75 \) for children with a behavioral disturbance and \( \text{rtt}=0.84 \) for normal children (Simons, 2014).
The Gibson Spiral Maze (GSM) has the Porteus test as precursor (Martin & Warde, 1971). It measures speed and precision in a paper and pencil test and consists of a spiral with a length of 235 cm with obstacles in the form of the letter O that are scattered along the pathway (Gibson, 1964). The individual has to draw a line from the inside of the spiral to the outside without touching the obstacles or the spiral. The test records the time required to fulfil the task and the number of errors that are made. These raw scores can be converted in percentiles and can be classified into four categories: slow and accurate, fast and accurate, slow and sloppy and fast and sloppy (Simons, 2014). In this study, the raw score for error and the time required to fulfil the task (in sec) were used. There was a negative relation found between speed and accuracy (r=0.42, p<0.001). The subject chooses for himself if speed or accuracy is the most important (Gibson, 1964). The interrater reliability for the Gibson Spiral Maze r=0.90–0.98 and the test-retest reliability for the time score is rtt=0.73 and for the error score rtt=0.77 (Gibson, 1964).

The Punktiertest für Kinder (PTK) (Schilling, 2009) is a test for children from 5 to 12 years old, designed by Schilling in 1974. This German test consists of a clown figure with 150 small circles on its circumference. During the test, the children have to put a dot inside these small circles. The variables that are measured are: the time required to fulfil the task and the errors that are made during the task. In this study, the raw score for error and the time required to fulfil the task (in sec) were used. The objectivity of the measurement is very high with r=0.99. The test-retest reliability for this test is r=0.92 and the validity is r=0.75 (Schilling, 2009).

The Movement-ABC test (M-ABC) is a frequently used assessment tool in The Netherlands in children from 4-12 years old to assess perceptual motor skills (Park & Simons, 2012). The test is divided in 4 age bands (4-6; 7-8; 9-10; 11-12 years), each with 8 items that are adapted to the specific age band and evaluate motor skills (Henderson & Sugden, 1992). In this study, the 4 age bands are used. The M-ABC tests three types of motor skills: “manual dexterity”
(3 items), “ball skills” (2 items) and “static and dynamic balance” (3 items). For every age band there are different items but they cover the same type of skills. Raw item scores are inverted to standard scores (0 to 5), a high standard score corresponds with a weak performance. The sum of all the standard scores is the total impairment score (TIS) with a maximum value of 40 that corresponds with poor perceptual motor skills. The TIS can be inverted to percentiles. A percentile of the TIS between the 5th and the 15th percentile indicates borderline motor difficulties. When the score is at the 5th percentile or less, the score indicates definite motor difficulties (Henderson, & Sugden, 1992). The sum of the standard scores for a specific type of motor skill can also be inverted to percentiles. In this study, the standard scores for the one subscale (manual dexterity) and the TIS were used. Reliability and validity of the M-ABC can be found in several studies (Wright, Sugden, Ng & Tan, 1994)

Procedure

The four tests were, where possible, done on one day by experienced researchers. The tests were administered in the mother tongue of the region (Dutch) and in random order. The randomization of the tests causes that confounding factors like fatigue, could have little influence on the results. Encouragement, during the tests, was given to augment the child’s effort. The testing time for the Epreuve de pointillage is one minute per hand. Only the results of the preferred hand were used in this study. The Gibson Spiral Maze takes 1-3 minutes to administer. The testing time of the PTK is 3-5 minutes per hand. Again only the results of the preferred hand were used in this study. The M-ABC test takes 20-30 minutes to administer (Henderson, & Sugden, 1992). From this test the standard scores for manual dexterity and the total impairment score were used.

Statistical analysis
Data were analysed by using Statistica software (version 12). First, the normality of the results of the entire test group was checked by using the Kolmogorov-Smirnov test. A Spearman correlation test was applied to check the associations between the FM assessment tools over the seven (7) years of the study. Where significant correlations were identified, the strength of the association was labelled according Portney and Watkins (2009). A Kruskall-Wallis Anova with post-hoc Kruskal-Wallis test was carried out to analyze age differences. In this study, not the chronological age but the educational level was used in the analysis of the age differences. The Kruskal-Wallis test was used to analyse gender differences. The level of statistical significance was set at \( p < 0.05 \).

Results

T-tests comparing the results of the children who were included into the study and children who were not included, indicated no significant differences on all the variables, meaning that the children of the experimental group were representative for the whole group of children.

Correlation between the variables of the FM assessment tools

The correlations carried out on the raw data, varied from .87 till .04. Most of the results were good to low (Portney & Watkins, 2009). The highest correlation was found between the results of the Epreuve de pointillage and the time score for the PTK (-.87). Especially the results of the correlation between the raw data of manual dexterity from the Mov-ABC with the other test results were low. All correlations are displayed in table 1.
Table 1. Spearman correlations between the test-results of The Epreuve de pointillage, the Gibson Spiral Maze, the Punktier Test für Kinder and the Movement-Assessment Battery for Children

<table>
<thead>
<tr>
<th></th>
<th>Epreuve de pointillage</th>
<th>Gibson Spiral Maze</th>
<th>Gibson Spiral Maze</th>
<th>Punktier- test</th>
<th>Punktier- test</th>
<th>Movement ABC</th>
<th>Movement ABC</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Time</td>
<td>Error</td>
<td>Time</td>
<td>Error</td>
<td>Manual Dexterity</td>
<td>Total</td>
<td></td>
</tr>
<tr>
<td>Epreuve de pointillage</td>
<td>.58*</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gibson Spiral Maze</td>
<td></td>
<td></td>
<td>Time</td>
<td>.39*</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Error</td>
<td></td>
<td></td>
<td></td>
<td>.04</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gibson Spiral Maze</td>
<td>.37*</td>
<td>.58*</td>
<td>.37*</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Punktier-test</td>
<td></td>
<td></td>
<td></td>
<td>-.46*</td>
<td></td>
<td>.45*</td>
<td></td>
</tr>
<tr>
<td>Time</td>
<td></td>
<td></td>
<td></td>
<td>.26*</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Error</td>
<td></td>
<td></td>
<td></td>
<td>.39*</td>
<td></td>
<td>.45*</td>
<td></td>
</tr>
<tr>
<td>Movement ABC</td>
<td>-.18*</td>
<td>.17*</td>
<td>.17*</td>
<td>.20*</td>
<td>.30*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Manual Dexterity</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Movement ABC</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>.32*</td>
<td>.33*</td>
</tr>
<tr>
<td>Total</td>
<td>-.06</td>
<td>.22*</td>
<td>.13</td>
<td>.32*</td>
<td>.33*</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Significant at 0.5% level

Ages differences

For educational level, a statistical significant effect (p<0.00005) was found for almost all variables except for the Movement ABC, but it was still significant (Table 2) meaning the results from preschool till the sixth grade varied over the time, but not in a systematic way. The post hoc analysis showed that significant differences were especially found between the results from the preschool and the results of the third grade and higher. This trend was observed in most of the variables (Table 3).
Table 2. Univariate results for educational level for The Epreuve de pointillage, the Gibson Spiral Maze, the PunktierTest für Kinder and the Movement-Assessment Battery for Children

<table>
<thead>
<tr>
<th></th>
<th>Preschool Mean (SD)</th>
<th>1st grade Mean (SD)</th>
<th>2nd grade Mean (SD)</th>
<th>3rd grade Mean (SD)</th>
<th>4th grade Mean (SD)</th>
<th>5th grade Mean (SD)</th>
<th>6th grade Mean (SD)</th>
<th>Kruskal-Wallis H</th>
<th>p-value</th>
<th>Kruskal-Wallis-test</th>
</tr>
</thead>
<tbody>
<tr>
<td>Epreuve de pointillage</td>
<td>58.82(12.35)</td>
<td>80.50(12.42)</td>
<td>83.64(12.05)</td>
<td>98.84(13.86)</td>
<td>107.14(15.81)</td>
<td>122.88(17.44)</td>
<td>128.82(19.36)</td>
<td>114.185</td>
<td>0.0000</td>
<td>0=3, 5, 6</td>
</tr>
<tr>
<td>Gibson Spiral Maze Time</td>
<td>85.27(51.35)</td>
<td>72.64(33.70)</td>
<td>82.82(18.39)</td>
<td>54.05(10.94)</td>
<td>58.41(11.78)</td>
<td>50.27(13.75)</td>
<td>32.86(11.18)</td>
<td>60.903</td>
<td>0.0000</td>
<td>0=3, 5, 6</td>
</tr>
<tr>
<td>Gibson Spiral Maze Error</td>
<td>35.32(29.10)</td>
<td>10.48(5.63)</td>
<td>9.95(6.83)</td>
<td>7.06(2.29)</td>
<td>3.68(3.07)</td>
<td>11.80(7.26)</td>
<td>3.64(2.36)</td>
<td>55.865</td>
<td>0.0000</td>
<td>0=3, 5, 6</td>
</tr>
<tr>
<td>Punktier Test Time</td>
<td>204.68(39.16)</td>
<td>137.50(22.63)</td>
<td>128.82(12.24)</td>
<td>107.06(14.50)</td>
<td>98.91(14.21)</td>
<td>88.06(13.80)</td>
<td>88.82(13.10)</td>
<td>112.569</td>
<td>0.0000</td>
<td>0=3, 5, 6</td>
</tr>
<tr>
<td>Punktier Test Error</td>
<td>95.50(29.10)</td>
<td>76.14(38.98)</td>
<td>91.27(9.98)</td>
<td>9.50(9.01)</td>
<td>22.90(20.20)</td>
<td>17.76(24.46)</td>
<td>17.28(34.60)</td>
<td>83.890</td>
<td>0.0000</td>
<td>0=3, 5, 6</td>
</tr>
<tr>
<td>Movement ABC Manual Score</td>
<td>4.00(2.85)</td>
<td>2.00(1.14)</td>
<td>3.16(1.99)</td>
<td>7.70(1.31)</td>
<td>4.25(2.78)</td>
<td>3.20(2.98)</td>
<td>2.61(2.45)</td>
<td>29.943</td>
<td>0.0004</td>
<td>0=3, 5, 6</td>
</tr>
<tr>
<td>Movement ABC Total</td>
<td>7.57(1.62)</td>
<td>4.43(1.32)</td>
<td>3.64(3.62)</td>
<td>1.77(1.75)</td>
<td>5.55(4.98)</td>
<td>7.16(5.90)</td>
<td>7.30(3.48)</td>
<td>27.262</td>
<td>0.0001</td>
<td>0=3, 5, 6</td>
</tr>
</tbody>
</table>

Remarkable is that the scores of most of the results in the third grade were not in line with the general tendency of the results, they are or better or worse. For the Movement ABC, the blocks that could be observed are in agreement with the age periods of the different scales of the Movement ABC.

The number of errors as well in the Gibson as in the PTK showed also a similar profile. They were getting down till the 4th grade and again from the 4th grade till the sixth grade.
Gender differences

Table 3. Univariate results for gender for The Epreuve de pointillage, the Gibson Spiral Maze, the PunktierTest für Kinder and the Movement-Assessment Battery for Children

<table>
<thead>
<tr>
<th>Test</th>
<th>Boys (n=10)</th>
<th>Girls (n=12)</th>
<th>Kruskal-Wallis test</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>SD</td>
<td>Mean</td>
<td>SD</td>
</tr>
<tr>
<td>Epreuve de pointillage</td>
<td>93.46</td>
<td>25.82</td>
<td>99.51</td>
<td>28.06</td>
</tr>
<tr>
<td>Gibson Spiral Maze</td>
<td>69.49</td>
<td>33.46</td>
<td>62.20</td>
<td>19.77</td>
</tr>
<tr>
<td>Error</td>
<td>11.33</td>
<td>14.03</td>
<td>12.49</td>
<td>16.88</td>
</tr>
<tr>
<td>PunktierTest</td>
<td>123.70</td>
<td>39.37</td>
<td>121.12</td>
<td>46.49</td>
</tr>
<tr>
<td>Error</td>
<td>42.71</td>
<td>44.45</td>
<td>30.86</td>
<td>38.01</td>
</tr>
<tr>
<td>Movement ABC Manual</td>
<td>3.92</td>
<td>2.88</td>
<td>2.11</td>
<td>2.42</td>
</tr>
<tr>
<td>Movement ABC Dexterity</td>
<td>7.03</td>
<td>5.28</td>
<td>4.90</td>
<td>4.68</td>
</tr>
</tbody>
</table>

For most of the results, significant differences were found between the results of boys and girls. Only for the PunktierTest the error score and the results for the Movement ABC were significant, girls doing better than boys. In general the boys were slower than and the girls had less errors, but this was non-consistent.

This article is distributed under the terms of the Creative Commons Attribution 4.0 License (https://creativecommons.org/licenses/by-nc-nd/4.0/) which permits to share — copy and redistribute the material in any medium or format.
Discussion

The main aim of the present study was to analyze the convergent validity between the four FM assessment tools. Generally, low correlations were identified between the four FM assessment tools. These generally low correlations might indicate that the assessment tools measure different FMS aspects. These differences between the four assessment tools might be explained by looking back to the taxonomy of Fleishman and Ellison (1962). The higher correlations between the PTK (time score) and the Epreuve de pointillage might be clarified by the fact that both assessment tools mainly measure the perceptual motor ability of aiming according to the taxonomy of Fleishman and Ellison (1962). In both assessment tools, the children have to perform a similar movement in which they have to dot a small circle (PTK) or square (Epreuve de pointillage). Aiming differs from the perceptual motor abilities that are measured with the GSM and the M-ABC. According to the taxonomy of Fleishman and Ellison (1962), the GSM mainly measures the wrist-finger speed. The correlations, with the M-ABC, were therefore mainly found with manual dexterity subscale and to a lesser extent with the total impairment score (that includes all three subscales). According to the taxonomy of Fleishman and Ellison (1962), the three items of the manual dexterity subscale each measure a different perceptual motor ability: finger dexterity, arm-hand steadiness and wrist-finger speed. Despite the fact that the correlations with the manual dexterity subscale were higher, the correlations between the three other assessment tools were generally highest. This may possibly have something to do with the fact that the score of the manual dexterity subscale is a combination of the three item scores. This might influence the correlation with the other assessment tools because each item measures a different perceptual motor ability. Another possible confounding factor in the combination of the scores might be that two out of three items contain a time score (seconds) and one out of three contains an error score (Henderson & Sugden, 1992).
A second aim was to evaluate the influence of age on FMS. It was hypothesized that there would be a positive evolution in FMS in function of age. The results of the four tests improved significantly with age. According to Smith et al. (2000), there is a linear increase in FMS in children across ages five to ten and a curvilinear relationship in children across ages ten to twenty years. These relationships were also found in the present study, in which the significant differences between the test scores of the educational levels were mainly found in the lower groups (preschool, first grade and second grade), as opposed to the higher groups (fourth, fifth and sixth grade). This might indicate that the positive evolution of FMS mainly takes place in the first grades of elementary school. Henderson and Sugden (1992) based the standardization of the M-ABC on the finding that older children are more proficient than younger children. Another explanation might be the fact that several age bands of the M-ABC were used. The four age bands (4-6, 7-8, 9-10, 11-12 years old) contain different tasks, so the children might have felt it more difficult to perform the tasks of the higher age bands as compared with the lower age bands. In this study, educational level instead of chronological age was used to evaluate the influence of age on FMS. In future similar longitudinal research, it might be useful to use the M-ABC second edition (Henderson et al., 2007), and more specifically the Dutch version of the M-ABC second edition (Smits-Engelsman et al., 2010). The differences between the traditional M-ABC and the M-ABC second edition (Henderson et al., 2007) are: an age extension (from 3 to 16 years), a reduction from 4 to 3 age bands, clearer test instructions and new items for the balance subscale. These 3 larger age bands could possibly be better to identify the age differences than the 4 smaller age bands of the traditional M-ABC. In 2010, a Dutch adaptation of the M-ABC second edition with population specific Dutch and Flemish norms was developed by Smits-Engelsman et al. (2010). This Dutch version could be very useful because the population of this study is similar to the population to which the norms of the assessment tool are based on.
Thirdly, it was hypothesized that the results of the four FM assessment tools would show no significant difference between boys and girls. Our data analysis revealed that the results of the girls were significantly better than the results of the boys in the test scores of the M-ABC (TIS and manual dexterity), and the error score of the PTK. In earlier studies, researchers found that girls outperform boys in manual dexterity in the M-ABC (Maeland, 1992; Sigmundsson & Rostoft, 2003; Ruiz, Graupera, Gutiérrez & Miyahara, 2003). As opposed to the M-ABC, there was to our knowledge no earlier research done to investigate gender influences in the three other assessment tools. The results indicate that girls achieve better results. This would possibly presume that girls have better FMS. In general the conclusion of Malina and Bouchard (1991) can be followed that girls are better in FM abilities.

The limitations of this study are the rather small sample size, the fact that all children were recruited in one school and the high socio-economic status of the children. These limitations might provoke a problem with the generalization of our findings. The children were recruited from one school for practical reasons, because the children were followed for several years. The effect of the socio-economic status on FMS has been documented in the literature (Henderson & Sugden, 1992; Piek, Dawson, Smith & Gasson, 2008; Engel-Yeger et al., 2010). This effect can possibly be elucidated by the fact that parents from a higher socio-economic level may be more aware of the importance of exposing their children to a variety of games and after curricular activities that focus on FM development. So, it was found that the socio-economic status forecast the FMS at school age (Piek et al., 2008). Since in this study most children are coming from a family with a higher socio-economic status, the confounding effect of this variable was ignored. Because of these limitations and the restricted possibility to extrapolate the findings of this study, further research is needed to confirm the findings of this study.
For this study we did choose for the original Movement ABC- version of 1992, despite the fact that in the past years, new versions of it were published with adapted norms to different countries (Barnett & Henderson, 1998). But this is in our opinion a weak point. The new version is still called Movement ABC’s, but the translators have made adaptations to different items. This includes an adapted scoring method as an adapted way of calculating the raw score. So, in our opinion the results of the different version are difficult to compare. As we did use the raw data of the test of Epreuve de pointillage, there is no problem to use the test also on 5 years old ones despite the fact the test has only norms from six years on.

The main conclusion of this study was that the results of the Epreuve de pointillage and the Punktiertest für Kinder had the highest correlations. Secondly, there can be observed that there were differences between boys and girls in FMS. A final conclusion of this study was the fact that FMS improved significantly with age, especially in the lower grades. More research is needed to draw clear conclusions on this topic.

Declaration of Interest: The authors report no declaration of interest.
References


This article is distributed under the terms of the Creative Commons Attribution 4.0 License (https://creativecommons.org/licenses/by-nc-nd/4.0/) which permits to share — copy and redistribute the material in any medium or format.


