



The development of the  
**ATHLETIC SKILLS TRACK**

*A new motor competence assessment*

**Joris Hoeboer**



**THE DEVELOPMENT OF THE  
ATHLETIC SKILLS TRACK  
A NEW MOTOR COMPETENCE ASSESSMENT**

## Colophon

ISBN: 978-94-6323-585-3  
Photo cover: Sanne de Vries  
Lay out: Ilse Modder, [www.ilsemodder.nl](http://www.ilsemodder.nl)  
Printed by: Gildeprint – Enschede



© Joris Hoeboer, Den Haag, 2019

All rights reserved. Illustrations and brief abstract from this thesis may be used for scientific and educational purpose provided that the source is acknowledged.

The research presented in this thesis was a collaboration of The Hague University of Applied Sciences and VU Amsterdam.

This PhD thesis was embedded within the Amsterdam Movement Sciences Research Institute, at the Department of Behavioural and Movement Sciences, VU Amsterdam, the Netherlands.

This work was financially supported by a grant. The author received a PhD-grant for teachers from The Dutch National Science Organization (NWO) (9023.006.005).

VRIJE UNIVERSITEIT

**THE DEVELOPMENT OF THE  
ATHLETIC SKILLS TRACK  
A NEW MOTOR COMPETENCE ASSESSMENT**

ACADEMISCH PROEFSCHRIFT

ter verkrijging van de graad Doctor of Philosophy  
aan de Vrije Universiteit Amsterdam,  
op gezag van de rector magnificus  
prof.dr. V. Subramaniam,  
in het openbaar te verdedigen  
ten overstaan van de promotiecommissie  
van de Faculteit der Gedrags- en Bewegingswetenschappen  
op dinsdag 25 juni 2019 om 13.45 uur  
in de aula van de universiteit,  
De Boelelaan 1105

door

**Joris Johannes Albertus Andreas Maria Hoeboer**  
geboren te Rozenburg

promotor:

prof.dr. G.J.P. Savelsbergh

copromotor:

dr. S.I. de Vries

*“The only true wisdom is in  
knowing you know nothing”*

*Socrates (469 v.C. - 399 v.C.)*

evaluation committee:

prof.dr. C. Schuengel  
prof.dr. M. Lenoir  
prof.dr. K. de Martelaer  
dr. R. Mombarg  
dr. A. Ledebt





# TABLE OF CONTENTS

<b>Chapter 1:</b>	<b>13</b>
General introduction	
<b>Chapter 2:</b>	<b>29</b>
Validity of an Athletic Skills Track among 6- to 12-year old children	
<b>Chapter 3:</b>	<b>53</b>
Reliability and concurrent validity of a motor skill competence test among 4- to 12-year old children	
<b>Chapter 4:</b>	<b>75</b>
The Athletic Skills Track: Age- and gender-related normative values of a motor skills test for 4- to 12-year-old children	
<b>Chapter 5:</b>	<b>95</b>
Children's enjoyment of a motor skill test in Physical Education	
<b>Chapter 6:</b>	<b>113</b>
Validity and feasibility of an obstacle course to assess fundamental movement skills in a pre-school setting	
<b>Chapter 7:</b>	<b>137</b>
Reproducibility, validity and feasibility of motor competence assessment instruments: a systematic review	
<b>Chapter 8:</b>	<b>161</b>
Epilogue	
<b>Addendum</b>	<b>179</b>
Dutch summary (samenvatting)	180
Acknowledgements	184
About the author	188
List of Publications	189
Publications as part of this thesis	189
Other publications	190
Scientific conference proceedings	192

Other Media	194
Awards	195
Appendix 1: Description of the Athletic Skills Tracks	196
Appendix 2: Test manual Athletic Skills Track	197
Appendix 3: Descriptive statistics of respondents' characteristics	208
Appendix 4: 10-item evaluation checklist	209
Appendix 5: Reference list for included studies	210
Appendix 6: Table reliability included studies	214
Appendix 7: Table validity included studies	216
Appendix 8: Evaluation of all included studies	219



# 1

General introduction

## 1 General introduction

An active lifestyle is related to all kinds of positive health effects (Janssen & LeBlanc, 2010; Pate et al., 1995), nevertheless there is a worldwide increase of childhood obesity and physical inactivity that has become a public health problem (WHO, 2016). Children's and adolescents' physical activity levels have decreased worldwide in the last decades (Hallal et al., 2012). In the Netherlands physical activity levels are also dropping (Van Mechelen, Twisk, Post, Snel, & Kemper, 2000). Motor competence (MC) levels appear to have minified as well over the last decades among Dutch youth (Inspectie van het onderwijs, 2018; Runhaar et al., 2010). The promotion of sustainable MC and physical activity and healthy body weight in children and adolescents has become a global challenge (Robinson et al., 2015).

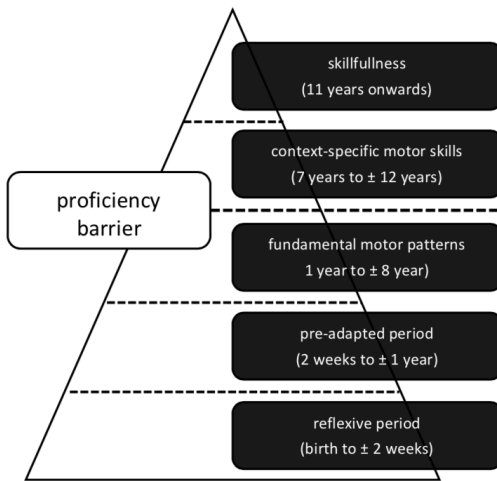
### 1.1 Physical literacy

In response to these before mentioned alarming trends a number of countries have placed more and more attention on physical education (PE), physical activity and sports promotion worldwide through the deployment of a concept called physical literacy (Giblin, Collins, & Button, 2014). This concept has gained interest in recent years as an important construct that may help to explain why people may or may not engage in physical activities and is being embraced in Great Britain, Canada, the United States and New Zealand (Whitehead, 2010). Physical literacy can be defined as the motivation, confidence, physical competence, knowledge and understanding to value and take responsibility for involvement in physical activities for life (Whitehead, 2013). The promotion of physical literacy is relevant throughout life and is therefore important for children as well as for adults and older adults.

MC is considered an important component of physical literacy (Giblin et al., 2014). Fundamental movement skills (FMS) play an important role in the development of MC and therefore can be seen as a way of developing the MC component of physical literacy (Almond, 2013). Physical literacy has been included as one of the focus areas of PE in the primary school setting in the United States. Researchers stated that PE is an important resource environment to develop physical literacy (Loprinzi, Davis, & Fu, 2015). Although physical literacy considers more than PE, the PE setting could provide children the foundations of physical literacy on which to build a lifelong engagement to, and enjoyment of, an active lifestyle (Edwards et al., 2017).

1.2 Definition and development of motor competence

Although numerous researches have been carried out in relation to MC, the term MC has become confusing because it has been used in various articles as a concept that is defined in different manners (Logan et al., 2018). In general MC can be seen as a term used to describe goal-directed human movement (Robinson et al., 2015). In this thesis MC will be defined as a person’s ability to execute different motor acts, including coordination of motor skills that are necessary to manage everyday tasks (Henderson & Sugden, 1992). This idea is based on the concept that coordination (e.g., coupling, spatial orientation, and balance ability) is essential in



**Figure 1.1** The mountain of Motor Development (Clarke & Metcalfe, 2002)

developing MC (Wormhoudt, Teunissen, & Savelsbergh, 2012). Motor coordination can be defined as the ability to have body segments work together in an organized manner (Turvey, 1990) and is considered an underlying component of FMS (Barnett et al., 2016). MC can be disentangled in a number of FMS (e.g., throwing, catching, hopping) that should be mastered during the preschool and primary school period (Branta, Haubenstricker, & Seefeldt, 1984; Gallahue & Ozmun, 2006). These FMS provide a foundation for children to develop

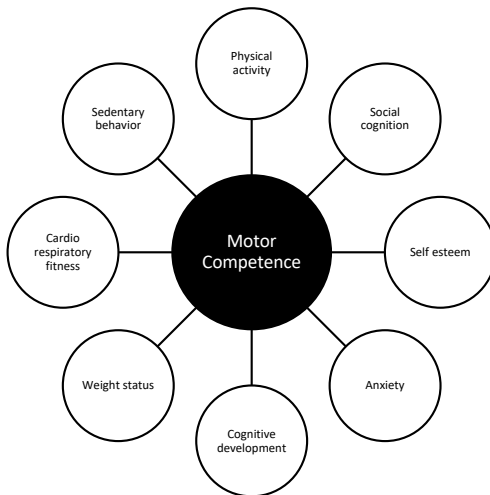
a more specialized movement repertoire, such as sport-specific movement skills for example; making a lay-up in basketball (Clarke & Metcalfe, 2002). FMS can be considered as the basis of lifelong movement skills such as swimming and cycling (Hulteen et al., 2015). FMS can be divided in three groups e.g.: object control/manipulative skills (i.e., throwing, catching, dribbling, kicking, strike, underhand roll), locomotor skills (i.e., walking, running, jumping, hopping, leaping, galloping, sliding, skipping), and balance/stability skills (i.e., body rolling, bending, dodging, stick balancing, one-foot balance, stretching, swinging, turning, and twisting) (Gallahue, Ozmun, & Goodway, 2012). Locomotor and balance skills are critical during the early childhood years (Robinson & Wadsworth, 2012; Cliff, Okely, Smith, & McKeen, 2009). This aligns with motor development literature, as locomotor skills develop earlier than object control/ manipulative skills (Robinson et al., 2015).

Seefeldt (1980) was one of the scientists who was interested in the relationship between learning FMS at an early age and the effect on physical activity on the long term. Seefeldt introduced the term “skills-barrier”. If a child does not reach a certain basic level (the “critical threshold”) in terms of FMS, the interest in physical activity will decrease drastically as the child gets older (see Figure 1.1). Children with FMS levels above this critical threshold are considered capable of applying their motor skills in life-long physical activities. But children with a skill level below this critical threshold will be less successful in physical activity, resulting in a higher risk of dropping out. Based on the work of Seefeldt (1980) the mountain of motor development was introduced by Clarke & Metcalfe (2002). In this conceptual model the crucial development of FMS is explained (see Figure 1.1). At individual, contextual or task level, a period or transitions from and to a period can differ. The metaphorical mountain is divided into five periods, starting with the reflexive period at the base. This is the first two weeks of life and includes reflexes such as squeezing, sucking and gagging. Very shortly after the reflexive period, the behaviour of the baby is more goal-oriented and spontaneous. This marks the beginning of the pre-adaptive period, in which the “species”-typical movements dominate. This period ends as independent walking and self-feeding, the beginning of the fundamental motor cycle period starts. This is characterized by the learning of the basic movements, or the FMS, which form the foundation for later higher motor skills. The fourth period is the context-specific motor skills period. In this period the basic movement patterns are adapted to specific, especially sports and game related purposes. The top of the mountain is the skill period. In this period an individual has 'grown' in terms of motor skills. Over time, an individual will have to adapt to changing situations, such as injuries, aging and other physical changes. This change has led to the introduction of the compensation period (Clarke & Metcalfe, 2002). However, this is not included in the standard mountain or pyramid model where each layer of the pyramid builds on the previous one. Clark & Metcalfe's mountain has helped to attract more interest in the effects of early motor learning. This paved the way for the conceptual models presented by Stodden et al. (2008) and later on Robinson et al. (2015).



### 1.3 Motor competence and health-related variables

To understand and turn around the disturbing trends described in paragraph 1.1, an increasing amount of research has been conducted on the relation between MC and health outcomes. Researchers state that the development of MC may be an important conducive factor for the development of positive weight status, health-related fitness and physical activity across childhood. Several recent cross-sectional studies (D’Hondt et al., 2011) and longitudinal



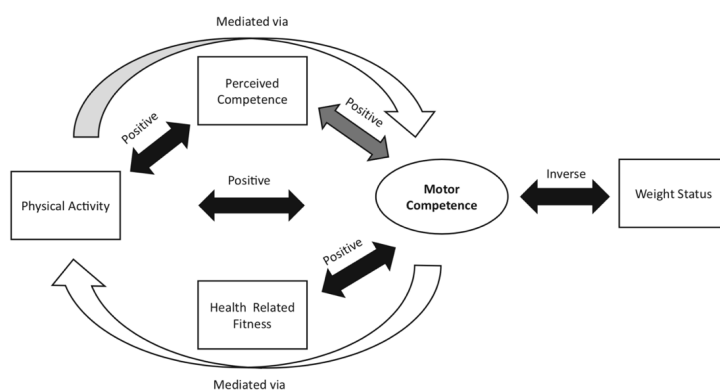
**Figure 1.2** Relations between Motor Competence and child’s health outcomes.

studies (Lopes et al., 2012a) demonstrated an inverse relationship between MC and body weight status in both sexes (Guo et al., 2018; Lopes et al. 2012b; D’Hondt et al., 2011). Physical activity levels are related with MC in cross-sectional (Logan et al., 2014; Lubens et al., 2010; Wrotniak et al., 2006) as well as in longitudinal studies (Lloyd et al., 2014). In both settings it showed that higher MC levels were related with higher physical activity levels. Higher MC levels are also related with higher health-related fitness such as

cardiorespiratory fitness (Okely, Booth, & Patterson, 2010). The pendant of MC is inversely related with PA, sedentary behaviour (Adank et al., 2018; Gu, 2016). Next to the health-related parameters that are affiliated to MC there has even been an increasing acknowledgment in the child development literature of the importance of MC in children’s cognitive development (Piek, Hands, & Licari, 2012). The development of MC is positively associated with cognitive development (Best, 2010; Hillman, Erickson, & Hatfield, 2017) and cognitive functioning (Chadlock et al, 2012., Hillman, Kamijo, & Scudder, 2011), as well as with social cognition and language skills (Leonard & Hill, 2014). Furthermore, MC performance is positively associated with self-esteem (Piek, Baynam, & Barrett, 2006) and reduced levels of anxiety (Skinner & Piek, 2001). Figure 1.2 shows the relations of MC with health-related parameters as well as with child’s general developmental aspects. Thus, MC is related to all kinds of health outcomes (Barnett et al., 2016).

Research into MC in relation to PA, health-related fitness and weight status has led to a conceptual model to provide insight about the relationships between these health-related outcomes and MC (Stodden et al., 2008) (see Figure 1.3). In this model, the relationships and how those relations strengthen each other are discussed. Based on this model, researchers have investigated those relationships even further. This has led to an update of the original model (Stodden et al., 2008) by Robinson et al. (2015). This model reflects the current consensus on the relationship between MC and health-related variables.

The model of Robinson et al. (2015) confirms that there is a positive relationship between MC and physical activity levels and health-related fitness. The model shows that MC and weight status have an inverse relationship within both childhood and adolescence. In addition, it becomes clear that the relationship between perceived MC and MC is already reasonably supported by research (Khodaverdi et al., 2016). Perceived MC has been identified as a potential mediator in the MC and physical activity levels pathway. The other mediator in the model, health related fitness as a mediator between MC and physical activity levels is yet to be confirmed empirically. The conceptual model is of great importance because it shows health-related variables and how they are related. How these factors affect health and how these factors change over time is a crucial focus for future research in this area. This knowledge might help to adjust the decreasing levels of MC, PA and health-related fitness and the increasing of weight status in childhood and adolescence (Robinson et al., 2015).



**Figure 1.3** Research consensus on motor competence and health-related variables. Black arrow indicates extensively tested: consistent relationship; dark grey arrow indicates moderately tested: variable relationship; partial grey arrow indicates partially tested: some evidence; white arrow indicates limited testing (Robinson et al., 2015).

Despite the importance, many cross-sectional and longitudinal studies involving relationships among variables in the model remains speculative without well-conducted experimental evidence. Intervention research with the ambition of targeting relationships in the model should be initiated during the early childhood years, as MC and PA levels behaviours should be established early in life and then tracked into the adult years.

#### 1.4 Measuring motor competence

To increase our understanding of the development of MC and the factors that can influence this development during childhood and factors that can influence this development, there is a need for large scale longitudinal data. Intervention studies to improve MC development related to the parameters that might influence MC development, and eventually health outcomes, seem to be limited in quantity and quality (Riethmuller, Jones, & Okely, 2009). Riethmuller, Jones, & Okely (2009) recommended that PE teachers should be involved in the implementation of such interventions because they have all children in their lessons. To be able to gather longitudinal data a reliable and valid assessment tool is needed that fits the PE setting. In recent decades, several MC assessments have been developed. Those assessments are not very feasible in the PE setting (Cools, De Martelaer, Samaey, & Andries, 2009). This is confirmed in a study among 1083 Dutch school principals (Reijgersberg & Lucassen, 2013). Most existing MC assessments are very time-consuming since it takes at least 20 minutes to measure one individual child (Cools et al., 2009). This makes it difficult to integrate the measurement in one single, one-hour, PE lesson. The assessments also require special test materials and extensive knowledge of the test protocols to be able to conduct the tests.

Cools et al. (2009) also suggest that further research in measuring FMS should involve PE teachers. It seems important to screen and monitor children's FMS over time to increase our understanding of the factors within the model of Robinson et al. (2015) (Cools et al., 2009; Lloyd et al., 2014; Stodden et al., 2009). The urgency to increase our understanding of motor skill development is also recognised by PE teachers who are willing to monitor motor skill competence of children more objectively (Lander, Morgan, Salmon, & Barnett, 2016). Despite the fact that PE teachers would like to monitor the motor skill competence of children more objectively, the mentioned shortcomings seem to be the reason many PE teachers currently do not use MC assessments structurally. Therefore, a reliable and valid assessment tool that fits the PE setting is needed.

### 1.5 An obstacle course based on the Athletic Skills Model

In 2012 the Athletic Skills Model (ASM) was introduced (Wormhoudt et al., 2012). The ASM finds its basis in scientific findings about MC development, talent development and experience from practitioners with these findings in practice. The goal of the ASM is to guide children to a active healthy life style in an optimal way by focussing on welfare, health and talent development of children and adolescents. Within the ASM, adaptability is essential and substantiated with coordinative abilities (CA), conditions of moment (COM) and FMS which are essential to create the possibility to a sporting healthy lifestyle (Wormhoudt et al., 2012). In the early days of the ASM a pilot study was performed to get more insight in measuring MC by developing an obstacle course based on the ASM. This obstacle course was developed by Wormhoudt et al. (2012) in cooperation with PE teachers to make sure the obstacle course fits the PE setting. This obstacle course was based on CA and FMS and was the foundation of the athletic skills track (AST) presented in this thesis.

### 1.6 Research aim

The purpose of this thesis is to examine the reliability, validity and practical usability of a new assessment tool: The Athletic Skills Track to assess fundamental movement skills among 4- to 12-year-old children in a physical education setting. The main research question can be formulated as follows:

How can physical education teachers assess motor competence to identify motor competence levels of 4-to 12-year old children during a regular physical education lesson in a reliable, valid and feasible manner?

To answer the main question, three sub-questions were formulated:

- 1) What is the reliability of an athletic skills track for measuring motor competence in children aged 4-12 years?
- 2) What is the validity of an athletic skills track for measuring motor competence in children aged 4-12 years?
- 3) What are reference and norm values for the athletic skills track for boys and girls aged 4-12 years?

## 1.7 Reading guide

In **chapter 2**, the feasibility and validity of the AST to assess fundamental movement skills among 6- to 12-year-old children in a physical education setting is presented. The AST discussed in this chapter is based on a pilot study that was performed to test if the AST offers opportunities for measuring FMS with time to complete the AST as solely outcome measurement. In the main study, the AST was refined and validated on a larger scale in a regular PE setting.

**Chapter 3** describes the investigation of the test-retest reliability, internal consistency and concurrent validity of three Athletic Skills Tracks that were developed based on the AST presented in chapter 2. Because of the ceiling effect that was found in the first version of the AST (see chapter 2), three age-related tracks were developed to measure MC of children aged 4- to 12-year old in the PE setting. In **chapter 4** age- and gender-related normative values belonging to the AST are presented. In a large-scale study a total of 7977 Dutch children, performed an age-related version of the AST as presented in chapter 3. The focus in **chapter 5** is children's enjoyment when being tested with the AST. Testing in PE can be valuable, but literature shows that it is important that the assessment is a positive experience for all children. **Chapter 6** is the result of a collaboration with the University of South Australia in Adelaide. This chapter describes the relationship between the AST and the test of gross motor development-II in a preschool setting. In **chapter 7** a systematic review about MC assessments is presented. In this chapter the reliability, validity and feasibility of MC assessments presented in scientific literature between 2000 and 2018 is discussed. A summary and discussion of the general findings are presented in the epilogue in **chapter 8**. In the epilogue the main findings, practical implication and suggestions for future research are presented.

## 1.8 References

Adank, A.M., Van Kann, D. H. H., Hoeboer, J. J., de Vries, S. I., Kremers, S. P. J., Vos, S. B. (2018). Investigating Motor Competence in Association with Sedentary Behavior and Physical Activity in 7- to 11-Year-Old Children. *International Journal of Environmental Research and Public Health* 15, 11: 2470.

Almond, L. (2013). Translating physical literacy into practical steps: the role of pedagogy. *ICSSPE Bull J Sport Sci Phys Educ*, 65:63–71.

Barnett, L. M., Lai, S. K., Veldman, S. L., Hardy, L. L., Cliff, D. P., Morgan, P. J., & Rush, E. (2016). Correlates of gross motor competence in children and adolescents: a systematic review and meta-analysis. *Sports Medicine*, 46(11), 1663-1688.

Best, J.R. (2010). Effects of physical activity on children's executive function: contributions of experimental research on aerobic exercise. *Developmental Review*; 30(4):331–351.

Branta, C., Haubenstricker, J., Seefeldt, V. (1984). Age changes in motor skills during childhood and adolescence. *Exercise and Sport Sciences Reviews*, 12(1):467–520.

Chaddock, L., Hillman, C. H., Pontifex, M. B., Johnson, C. R., Raine, L. B., & Kramer, A. F. (2012). Childhood aerobic fitness predicts cognitive performance one year later. *Journal of Sports Sciences*, 30(5), 421-430.

Clark, J. E., & Metcalfe, J. S. (2002). The mountain of motor development: A metaphor. *Motor development: Research and rReviews*, 2, 163-190.

Cliff, D. P., Okely, A. D., Smith, L. M., & McKeen, K. (2009). Relationships between fundamental movement skills and objectively measured physical activity in preschool children. *Pediatric Exercise Science*, 21(4), 436-449.

Cools, W., De Martelaer, K., Samaey, C., & Andries, C. (2009). Movement skill assessment of typically developing preschool children: A review of seven movement skill assessment tools. *Journal of Sports Science & Medicine*, 8(2), 154-68.

D'Hondt, E., Deforche, B., Vaeyens, R., Vandorpe, B., Vandendriessche, J., Pion, J., Lenoir, M. (2011). Gross motor coordination in relation to weight status and age in 5-to 12-year-old boys and girls: A cross-sectional study. *International Journal of Pediatric Obesity*, 6(sup3), e556–e564. doi:10.3109/17477166.2010.500388

Edwards, L. C., Bryant, A. S., Keegan, R. J., Morgan, K., & Jones, A. M. (2017). Definitions, foundations and associations of physical literacy: a systematic review. *Sports Medicine*, 47(1), 113-126.

Gallahue, D. L., Ozmun, J. C., & Goodway, J. (2006). *Understanding motor development: Infants, children, adolescents, adults*. (pp. 248-270). Boston: McGraw-hill.

Giblin, S., Collins, D., & Button, C. (2014). Physical literacy: importance, assessment and future directions. *Sports Medicine*, 44(9), 1177-1184.

Gu, X. (2016). Fundamental motor skill, physical activity, and sedentary behavior in socioeconomically disadvantaged kindergarteners. *Psychology, Health & Medicine*, 21(7), 871-881.

Guo, H., Schenkelberg, M. A., O'Neill, J. R., Dowda, M., & Pate, R. R. (2018). How Does the Relationship Between Motor Skill Performance and Body Mass Index Impact Physical Activity in Preschool Children? *Pediatric Exercise Science*, 30(2), 266-272.

Hallal, P. C., Andersen, L. B., Bull, F. C., Guthold, R., Haskell, W., Ekelund, U., & Lancet Physical Activity Series Working Group (2012). Global physical activity levels: surveillance progress, pitfalls, and prospects. *The Lancet*, 380(9838), 247-257.

Henderson, S, Sugden, D. (1992). *Movement assessment battery for children*. London: The Psychological Corporation.

Hillman, C. H., Kamijo, K., & Scudder, M. (2011). A review of chronic and acute physical activity participation on neuroelectric measures of brain health and cognition during childhood. *Preventive Medicine*, 52, S21-S28.

Hillman, C. H., Erickson, K. I., & Hatfield, B. D. (2017). Run for your life! Childhood physical activity effects on brain and cognition. *Kinesiology Review*, 6(1), 12-21.

Hulteen, R. M., Lander, N. J., Morgan, P. J., Barnett, L. M., Robertson, S. J., & Lubans, D. R. (2015). Validity and reliability of field-based measures for assessing movement skill competency in lifelong physical activities: a systematic review. *Sports Medicine*, 45(10), 1443-1454.

Inspectie van het onderwijs (2018). *De staat van het onderwijs: Onderwijsverslag 2016–2017* [The state of education: Year report 2016–2017]. Utrecht: Inspectie van het Onderwijs.

Janssen, I., & LeBlanc, A. G. (2010). Systematic review of the health benefits of physical activity and fitness in school-aged children and youth. *The International Journal of Behavioral Nutrition and Physical Activity*, 7(1), 1– 16. doi:10.1186/1479-5868-7-40

Khodaverdi, Z., Bahram, A., Stodden, D., & Kazemnejad, A. (2016). The relationship between actual motor competence and physical activity in children: mediating roles of perceived motor competence and health-related physical fitness. *Journal of Sports Sciences*, 34(16), 1523-1529.

Lander, N., Morgan, P. J., Salmon, J., & Barnett, L. M. (2016). Teachers' perceptions of a fundamental movement skill (FMS) assessment battery in a school setting. *Measurement in Physical Education and Exercise Science*, 20(1), 50-62.

Lenoir, M., Vandorpe, B., D'Hondt, E., Pion, J., Vandendriessche, J., Vaeyens, R., & Philippaerts, R. M. (2014). *KTK-NL Körperkoordinationstest für Kinder*. Herwerkte, gehernormeerde en vertaalde uitgave van de KTK voor het Nederlandstalig gebied (Vol. 1, pp. 80). Destelbergen: Universiteit Gent.

Leonard, H. C., & Hill, E. L. (2014). The impact of motor development on typical and atypical social cognition and language: A systematic review. *Child and Adolescent Mental Health*, 19(3), 163-170.

Logan, S. W., Ross, S. M., Chee, K., Stodden, D. F., & Robinson, L. E. (2018). Fundamental motor skills: A systematic review of terminology. *Journal of Sports Sciences*, 36(7), 781-796.

Logan, S. W., Robinson, L. E., Getchell, N., Webster, E. K., Liang, L. Y., & Golden, D. (2014). Relationship between motor competence and physical activity: A systematic review. *Research Quarterly for Exercise and Sport*, 85(S1), A14.



Lopes, V. P., Maia, J. A., Rodrigues, L. P., & Malina, R. (2012a). Motor coordination, physical activity and fitness as predictors of longitudinal change in adiposity during childhood. *European Journal of Sport Science*, 12(4), 384-391.

Lopes, V. P., Stodden, D. F., Bianchi, M. M., Maia, J. A., & Rodrigues, L. P. (2012). Correlation between BMI and motor coordination in children. *Journal of Science and Medicine in Sport*, 15(1), 38-43.

Loprinzi, P. D., Davis, R. E., & Fu, Y. C. (2015). Early motor skill competence as a mediator of child and adult physical activity. *Preventive Medicine Reports*, 2, 833-838.

Lloyd, M., Saunders, T. J., Bremer, E., & Tremblay, M. S. (2014). Long-term importance of fundamental motor skills: A 20-year follow-up study. *Adapted Physical Activity Quarterly*, 31(1), 67-78.

Lubans, D. R., Morgan, P. J., Cliff, D. P., Barnett, L. M., & Okely, A. D. (2010). Fundamental movement skills in children and adolescents. *Sports Medicine*, 40(12), 1019-1035.

Van Mechelen, W., Twisk, J. W., Post, G. B., Snel, J., & Kemper, H. C. (2000). Physical activity of young people: The amsterdam longitudinal growth and health study. *Medicine & Science in Sports & Exercise*, 32(9), 1610–1616. doi:10.1097/00005768-200009000- 00014

Newell, K.M., (1986) Constraints on the development of coordination, in *Motor Development in Children: Aspects of Coordination and Control*, Wade MG, Whiting HTA, editors, The Netherlands, Martinus Nijhoff, Dordrecht, p. 341–360 [https:// doi.org/10.1007/978-94-009-4460-2 19](https://doi.org/10.1007/978-94-009-4460-2_19).

Okely, A. D., Booth, M. L., & Patterson, J. W. (2001). Relationship of cardiorespiratory endurance to fundamental movement skill proficiency among adolescents. *Pediatric Exercise Science*, 13(4), 380-391.

Pate, R. R., Pratt, M., Blair, S. N., Haskell, W. L., Macera, C. A., Bouchard, C., King, A. C. (1995). Physical activity and public health: A recommendation from the centers for disease control and prevention and the american college of sports medicine. 273(5), 402–407. doi:10.1001/jama.1995.03520290054029

Piek, J. P., Baynam, G. B., & Barrett, N. C. (2006). The relationship between fine and gross motor ability, self-perceptions and self-worth in children and adolescents. *Human Movement Science, 25*(1), 65-75.

Piek J. P., Hands B., & Licari, M. K. (2012). Assessment of motor functioning in the preschool period. *Neuropsychology Review, 22*(4), 402-413.

Reijgersberg, N, Lucassen, J. (2013) Nulmeting bewegingsonderwijs: onderzoek naar de organisatie van het bewegingsonderwijs in het primair onderwijs, Mulier Instituut.

Riethmuller, A. M., Jones, R. A., & Okely, A. D. (2009). Efficacy of interventions to improve motor development in young children: a systematic review. *Pediatrics, 124*(4), e782-e792.

Robinson, L. E., Stodden, D. F., Barnett, L. M., Lopes, V. P., Logan, S. W., Rodrigues, L. P., & D'Hondt, E. (2015). Motor competence and its effect on positive developmental trajectories of health. *Sports Medicine, 45*(9), 1273-1284.

Robinson, L. E., Wadsworth, D. D., & Peoples, C. M. (2012). Correlates of school-day physical activity in preschool students. *Research Quarterly for Exercise and Sport, 83*(1), 20-26.

Runhaar, J., Collard, D. C., Singh, A. S., Kemper, H. C., Van Mechelen, W., & Chinapaw, M. (2010). Motor fitness in dutch youth: Differences over a 26- year period (1980–2006). *Journal of Science and Medicine in Sport/Sports Medicine Australia, 13*(3), 323–328. doi:10.1016/j.jsams.2009.04.006

Seefeldt, V.D. (1980). Developmental motor patterns: implications for elementary school physical education. In: Nadeau CH et al. (red.), *Psychology of motor behavior and sport*, 314-323. Champaign (Ill): Human Kinetics.

Skinner, R. A., & Piek, J. P. (2001). Psychosocial implications of poor motor coordination in children and adolescents. *Human Movement Science, 20*(1-2), 73-94.

Stodden, D. F., Goodway, J. D., Langendorfer, S. J., Robertson, M. A., Rudisill, M. E., Garcia, C., & Garcia, L. E. (2008). A developmental perspective on the role of motor skill competence in physical activity: An emergent relationship. *Quest, 60*(2), 290-306.

Tidén, A., Lundqvist, C., & Nyberg, M. (2015). Development and initial validation of the NyTid test: A movement assessment tool for compulsory school pupils. *Measurement in Physical Education and Exercise Science*, 19(1), 34–43. doi:10.1080/1091367X.2014.975228

Turvey, M. T. (1990). Coordination. *American Psychologist*, 45(8), 938. doi:10.1037/0003-066X.45.8.938

Whitehead, M. (Ed.). (2010). *Physical literacy: Throughout the lifecourse*. Routledge.

Whitehead, M. (2013). Definition of physical literacy and clarification of related issues. *ICSSPE Bulletin*, 65(1.2).

W.H.O., (2016). *Final Report of the Commission on Ending Childhood Obesity*. Geneva: World Health Organization.

Wormhoudt R., Teunissen J. W., & Savelsbergh G. J. (2012). *Athletic skills model*. Nieuwegein: Arko Sports Media.

Wrotniak, B. H., Epstein, L. H., Dorn, J. M., Jones, K. E., & Kondilis, V. A. (2006). The relationship between motor proficiency and physical activity in children. *Pediatrics*, 118(6), e1758-e1765.



# Validity of an Athletic Skills Track among 6- to 12-year old children

Hoeboer, J., De Vries, S., Krijger-Hombergen, M., Wormhoudt, R., Drent, A., Krabben, K., & Savelsbergh, G. (2016).

*Journal of Sports Sciences, 34(21), 2095-2105*

The purpose of this study was to examine the feasibility and validity of an Athletic Skills Track (AST) to assess fundamental movement skills among 6- to 12-year-old children in a physical education setting. Four hundred sixty-three Dutch children (211 girls, 252 boys) completed three tests: the Körperkoordinationstest für Kinder (KTK) and two Athletic Skills Tracks (AST-1, AST-2). The validity of AST-1 and AST-2 was examined by correlating the time (s) needed to complete the tracks and the KTK Motor Quotient (MQ).

Overall, there was a low correlation between AST-1 and the KTK MQ ( $r = -0.474$  ( $P < 0.01$ )) and a moderate correlation between AST-2 and the KTK MQ ( $r = -0.502$  ( $P < 0.01$ )). When split up by age group the associations were much higher and ranged between  $r = -0.469$  and  $r = -0.767$ , with the exception of the low correlation coefficient of the AST-2 in 7-year-olds. The results indicate that fundamental movement skills of 6- to 12-year-old children can be assessed with a quick, convenient and low-cost motor competence test in a physical education setting, i.e., an Athletic Skills Track. Future studies should further assess the reliability, discriminative ability and validity of age-specific versions of the AST.

## 2 Validity of an Athletic Skills Track among 6- to 12-year old children

### 2.1. Introduction

Despite the well-known health effects of an active lifestyle, physical activity ratings among children and adolescents seem to be decreasing in many countries (Dollman, Norton, & Norton, 2005; Hallal et al., 2012; Salmon & Timperio, 2007). In the Netherlands, the percentage of children under 17 years of age who reach the public health guideline to accumulate a minimum of 60 min of moderate to vigorous physical activity per day is 18% (Hildebrandt, Bernaards, & Stubbe, 2013). According to Wrotniak, Epstein, Dorn, Jones, and Kondilis (2006) motor skills competence is a crucial component for children's engagement in physical activity. In line with children's physical activity level, motor skills appear to have dropped significantly as well over the last decades among Dutch youth (Runhaar et al., 2010), a trend that is also seen among Flemish youth (Vandorpe et al., 2011).

Therefore, stimulating physical activity has become a public health priority. Physical education (PE) is suggested to be an important component in efforts to promote physical activity (Control & Prevention, 2011). The PE setting has the necessary resources and provides access to all youth (Hardy, King, Espinel, Cosgrove, & Bauman, 2013). One of the goals of PE is to increase children's movement repertoire by supporting the development of motor skills. Despite this goal, measurement of children's motor skills development is rare in The Netherlands (Reijgersberg, van der Werff, & Lucassen, 2013).

Children's motor skills are positively correlated with physical activity (Bouffard, Watkinson, Thompson, Causgrove Dunn, & Romanow, 1996; Lubans, Morgan, Cliff, Barnett, & Okely, 2010; Okely, Booth, & Patterson, 2001; Wrotniak et al., 2006) and health-related physical fitness (Cantell, Crawford, & Doyle-Baker, 2008; Haga, 2008; Lubans et al., 2010; Stodden et al., 2008) in a cross-sectional research setting. There is some evidence that children's motor skills are also related to people's physical activity level on the long term. Lloyd, Saunders, Bremer, and Tremblay (2014) found a long-term relation between motor skills proficiency at age 6 and leisure time physical activity at age 26. This is in line with the conceptual models presented by Clark & Metcalfe and Stodden et al. (Clark & Metcalfe, 2002; Stodden et al., 2008; Stodden, Langendorfer, & Robertson, 2009). They hypothesise a relationship between physical activity and motor skill competence and state that Fundamental Movement Skill (FMS) competence is critical in encouraging a physically active lifestyle. To date, high-quality, large-scale, longitudinal studies on this topic are scarce (Cliff, Okely, Smith, and McKeen 2009; Lloyd et al., 2014). In order to increase our knowledge of the relationship between motor

skills at young age and physical activity later in life valid assessment tools are needed. These tools should be able to identify children with low motor coordination as early as possible, since the foundations of FMS are laid in early childhood (Lopes, Stodden, Bianchi, Maia, & Rodrigues, 2012; Runhaar et al., 2010).

FMS can be examined with several assessment tools. The most frequently used assessment tools in early childhood are: Motoriktest für Vier- bis Sechsjährige Kinder (MOT 4-6), Movement Assessment Battery for Children (Movement-ABC), Peabody Development Motor Scales (PDMS), Körperkoordinationstest für Kinder (KTK), Test of Gross Motor Development (TGMD), the Maastrichtse Motoriek Test (MMT) and the Bruininks-Oseretsky Test of Motor Proficiency (BOTMP) (Cools, Martelaer, Samaey, & Andries, 2009). Cools et al. (2009) reviewed these different assessment tools and concluded that in general the internal consistency and inter-rater reliability of those tools are high. In addition, the concurrent validity of most assessment tools is moderate. However, most tests are not very feasible in a PE setting. This is confirmed in a study among 1083 Dutch school principals (Reijgersberg et al., 2013). It takes at least 20 min to measure one individual child. Furthermore, special test materials and extensive knowledge of the test protocols are required to be able to conduct the tests. Purchase price of the tests range between €175 and €1375. In summary, Cools et al. suggest that further research in measuring FMS should involve PE teachers. It seems important to screen and monitor children's FMS over time with reliable, valid and feasible assessment tools (Cools et al., 2009; Lloyd et al., 2014; Stodden et al., 2009). In addition to the clinimetric properties of an assessment tool, the experience of children when being tested should also be taken into account. Comparing children with standards can have a negative impact on the motivation of those most in need of encouragement (Cale et al., 2014). Cale et al. (2014) have formulated a number of recommendations that PE teachers should take into account in order to weaken this negative impact as much as possible. One of the recommendations is to test within the context of a regular PE lesson. In addition, they indicate that it is important that the test is a positive experience for the children.

Existing FMS assessment tools do not meet all these criteria. Therefore, a new screening tool, i.e., an Athletic Skills Track (AST), was developed by Wormhoudt, Teunissen, and Savelsbergh (2012) in cooperation with PE teachers. The AST is grounded in existing theories of children's movement development (Clark & Metcalfe, 2002; Wormhoudt et al., 2012) and the crucial role of FMS in these theories. Clark and Metcalfe introduced the Mountain of Motor Development as a metaphor. In this "mountain" four phases are described; the first phase is the reflex phase (e.g., grasping and tonic neck reflexes), followed by the rudimentary phase (e.g.,

rolling, creeping and walking). Internalization of the FMS takes place in the movement pattern phase. The FMS can be divided into the following categories: (1) locomotive skills (e.g., walking, crawling and jumping), (2) manipulative skills (e.g., throwing and kicking) and (3) stability skills (e.g., dynamic and static balance and axial movements like rolling). In the last phase of the Mountain of Motor Development, i.e., the growth and refinement phase, movement becomes more complex. The FMS are considered to be a prerequisite for the last phase (Clark & Metcalfe, 2002). With the help of PE teachers all three FMS categories (locomotive, manipulative and stability skills) were translated into a track that consisted of a sequence of 10 exercises that are feasible in a PE setting. Children are asked to complete the track as quickly as possible. The time to complete the track is the only parameter that is measured. With this approach the AST is a fundamentally different tool than conventional FMS assessment tools. The new tool took into account the disadvantages of existing assessment tools (i.e., high cost, time-consuming, not suitable for a PE setting) and aimed to assess motor competence in general among large groups of school-aged children in a PE setting. Ideally, the new tool can be used for (1) screening: from identifying individuals at risk to talent identification; (2) monitoring: monitoring of motor development of individuals and monitoring trends in motor skill at (sub)group and school level over a longer period; (3) benchmarking: comparing groups and schools on the motor skills of children; and (4) evaluation: the evaluation of interventions (methods, programmes, products) to improve the motor skills of children.

To our knowledge, no studies have assessed children's motor competence with a skills track in a PE setting by measuring only the time to complete a track. Thus, the purpose of this study was to examine the concurrent validity and feasibility of the AST to assess motor competence in general among 6- to 12-year-old children in a PE setting.

We hypothesise that by measuring time (a quantitative measurement) to complete the AST it is possible to find a moderate association with a qualitative measurement of children's motor skills (Hinkle, Wiersma, & Jurs, 2003). Furthermore, it is assumed that the AST is feasible in the PE setting (less time consuming and based on regular PE equipment).



## 2.1 Methods

First, a pilot study was performed to test if the AST offers opportunities for measuring FMS in a PE setting. Next, in the main study, the AST was refined and validated on a larger scale.

## 2.2 Pilot study

In the pilot study a convenience sample of 54 children (28 boys, 26 girls) aged 6–12 years participated, all of them attending the same Dutch primary school located in the centre of Amsterdam, the Netherlands. The children completed the AST and a valid and reliable motor competence test, the KTK (Kiphard & Schilling, 2007). The AST consisted of a series of locomotive, manipulative and stability skills that had to be completed as fast as possible (e.g., barefoot alligator crawl for 5 m on a mat, travelling jumps through five rings, throwing and catching a ball, forward roll and clambering over a vaulting box).

The pilot study showed that the AST is a feasible test to assess FMS among children aged 6–12 years in a PE setting. It was possible to measure 24 children in a PE lesson of 50 min with the equipment that is available in every PE setting in The Netherlands. The overall motor quotient scores on the KTK (KTK MQ) showed a moderate correlation with AST scores (i.e., time to complete the track in seconds) corrected for age ( $r = -0.645$ ,  $P < 0.001$ ). The results of the pilot test suggest that the AST is a feasible and a valid tool to determine children's general motor competence. The aim of the validation study was to repeat the pilot study in a different context, with a larger study population and with the help of several research assistants. Therefore, the AST test protocol and the track were refined and standardised in order to gather reliable data in different contexts. For example, cones were added in the track to make sure that all the children followed the exact same route.

## 2.3 Validation study

### 2.3.1 Subjects

Children aged 6–12 years were recruited from five primary schools in the The Hague region, the Netherlands. The schools were selected at random from a database of the The Hague University of Applied Sciences for internship schools. Informed consent was obtained from the parents or guardians of 623 children after they were given written information about the purpose and nature of the study. The study protocol was approved by the Medical Ethical Committee of the Faculty of Human Movement Sciences, VU University Amsterdam, The Netherlands (ECB 2015–31).

### 2.3.2 Measurements

Measurements included three tests: the KTK and two ASTs (AST-1, AST-2). All measurements were conducted in a separate section of the gym by a team of four research assistants (fourth year PE students of the The Hague University of Applied Sciences) during regular PE lessons. One of the assistants organised and administrated the AST tracks. The other three assistants were responsible for the KTK. They had been trained in conducting the tests according to the protocols in four meetings. The protocols provided guidelines about how to deal with adverse events, such as falling. In practice, adverse events were absent. Testing was spread out over a 2-week time period in May 2014. To measure all the children in this period a uniform schedule was developed: during the first week the children were measured on AST-1, during the second week on AST-2. In this 2-week time period all children also completed the KTK. Before and during the testing period the children received PE lessons as planned.

#### 2.3.2.1 *Körperkoordinationstest für Kinder*

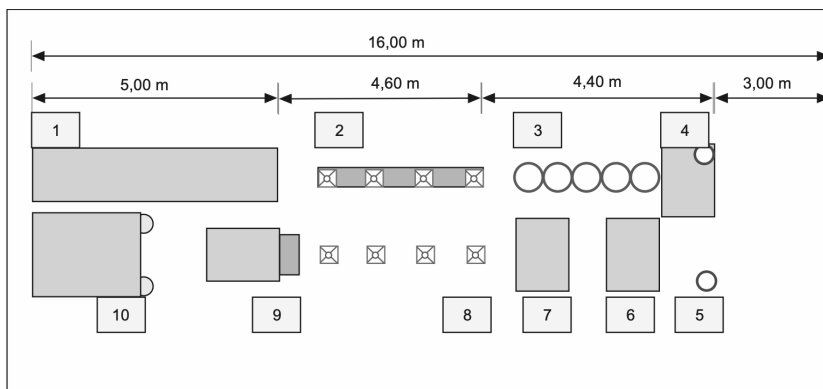
The KTK, developed and validated among German children (Kiphard & Schilling, 2007), was used as a reference measure to examine the concurrent validity of AST-1 and AST-2. The KTK is divided into four subtests;

- (1) Walking backwards three times along each of three balance beams (3 m length; 6, 4.5 and 3 cm width, respectively; 5 cm height).
- (2) Moving across the floor in 20 s by stepping from one plate (25 cm × 25 cm × 5.7 cm) to the next, transferring the first plate, stepping on it, etc.
- (3) Jumping from one leg over an increasing pile of pillows (60 cm × 20 cm × 5 cm each) after a short run-up.
- (4) Jumping laterally as many times as possible over a wooden slat (60 cm × 4 cm × 2 cm) in 15 s.

The test protocol for the Dutch language area (Lenoir et al., 2014; Vandorpe et al., 2011) was followed. Research assistants carried out the test protocol. It took ± 25 min to complete the KTK per child.

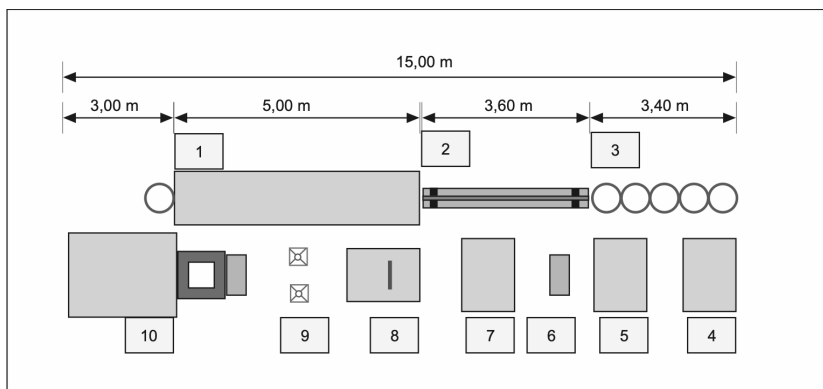
2.3.2.2 Athletic Skills Tracks

Based on the pilot study, two tracks were designed: AST-1 and AST-2. The tracks consisted of a series of fundamental motor tasks (n = 10) to be completed as fast as possible. AST-1 (see also Figure 2.1) consisted of the following locomotive, manipulative and stability skills: (1) alligator crawl, (2) bunny hops, (3) travelling jumps, (4) throwing and catching a ball, (5) kicking and stopping a ball, (6) forward roll, (7) backward roll, (8) running backwards, (9) clambering and (10) jumping.



**Figure 2.1** AST-1 schematically displayed. (1) crawling, (2) hopping, (3) jumping, (4) throwing and catching a ball, (5) kicking and stopping a ball, (6) front roll, (7) back roll, (8) walking backwards, (9) clambering and (10) jumping.

AST-2 (see Figure 2.2) consisted of the following locomotive and stability skills: (1) alligator crawl, (2) walking forward, (3) travelling jumps, (4) forward roll, (5) backward roll, (6) flank vault, (7) pencil roll, (8) tumbling forward, (9) hopscotch and (10) straight jump. The skills are further described in Appendix 1.



**Figure 2.2** AST-2 schematically displayed. (1) crawling, (2) balancing, (3) jumping, (4) front roll, (5) back roll, (6) hopping, (7) longitudinal roll, (8) tumble over on the rings, (9) hopscotch and (10) stretch jump.

All children were shown an instruction movie before they performed three try-out trials per track and two trials per track. During the three try-out trials the children received feedback from the research assistant if necessary.

A research assistant measured the time to complete the two trials with a stopwatch. After the first trial the children had 4–5 min rest before they performed the second trial.

### 2.3.3 Data analysis

Of the 623 children who were allowed by their parents or guardians to participate in the study, 463 children met the inclusion criteria (age: between 6 and 12 years; all three tests completed). First, the raw test scores on the KTK were converted into age- and gender-specific motor quotients (KTK MQ) using the test protocol for the Dutch language area (Lenoir et al., 2014). Participants were then classified into five categories based on the classification of Kiphard and Schilling (2007). Children with a MQ value between 86 and 115 are considered as having a normal gross motor coordination (NMG), between 71 and 85 as having a moderate gross motor coordination disorder (MMD) and children scoring 70 or less as having a severe gross motor coordination disorder (SMD). Children scoring between 116 and 130 are considered as having a good motor coordination (GMG) and children scoring 131 and higher as high motor gifted (HMG).

Second, descriptive statistics were performed to characterise the sample. To analyse a normal distribution of the outcome parameters, histograms were plotted and the kurtosis and skewness values for each of the outcome parameters were assessed. Differences between boys and girls were tested using an independent samples t-test. In order to examine a potential learning effect between the two trials of the ASTs the intraclass correlation coefficient (ICC) was calculated with a two-way random model. In addition, the scores on separate trials were compared using a paired samples t-test. To examine the concurrent validity of the tracks, KTK MQ and AST-1 and AST-2 scores were correlated with a Pearson's correlation test. The following rules of thumb have been used to interpret the size of the correlation coefficients: negligible: <0.30; low: 0.30–0.50; moderate: 0.50–0.70; high: 0.70–0.90; very high: 0.90–1.00 (Hinkle et al., 2003).

To analyse the distinctiveness of the ASTs to identify children with low or high motor skills a one-way ANOVA was performed comparing the scores on AST-1 and AST-2 between the five groups based on the KTK MQ. All statistical analyses were performed using IBM SPSS 22.0 64-bit edition. Values were considered statistically significant at  $P < 0.05$ .

## 2.4 Results

In total, 463 children (252 boys and 211 girls) with a mean age of  $9 \pm 2$  years completed all three tests. Data were normally distributed. The anthropometric characteristics of the children and the KTK MQ and AST scores are shown in Table 2.1 and 2.2. On average, it took about  $44 \text{ s} \pm 11$  to complete AST-1 and  $45 \text{ s} \pm 12$  to complete AST-2 per child. The average KTK MQ score was  $104 \pm 14$ . An independent-samples t-test showed that girls had a significantly higher KTK MQ than boys ( $106 \pm 13.3$  versus  $102 \pm 14.2$ ;  $t(461) = -3.242$ ,  $P < 0.01$ ). On the other hand, boys were significantly faster on AST-1 than girls ( $42 \pm 9.5 \text{ s}$  versus  $46 \pm 11.5 \text{ s}$ ;  $t(460) = -3.823$ ,  $P < 0.01$ ). When looking into AST-1/ KTK level in Table 2.2 it shows that 69.5% of the children were normally motor gifted. They completed AST-1 in  $52 \pm 14 \text{ s}$ . These children completed AST-2 in less time, i.e.,  $45 \pm 10 \text{ s}$ . Table 2.3 shows the results on the KTK, AST-1 and AST-2 per age group.

**Table 2.1** Descriptive statistics for each group of participants.

		N	Mean	( $\pm$ SD)	95% CI	
					lower	upper
Age (years)	<b>Total</b>	463	9	(2)	8.95	9.27
	<b>Boys</b>	252	9	(2)	8.88	9.33
	<b>Girls</b>	211	9	(2)	8.89	9.36
Length (cm)	<b>Total</b>	459	138.7	(12.1)	137.6	139.9
	<b>Boys</b>	250	138.4	(11.0)	136.9	139.7
	<b>Girls</b>	209	139.1	(13.4)	137.2	140.9
Weight (kg)	<b>Total</b>	459	33.7	(9.1)	32.9	34.6
	<b>Boys</b>	250	33.8	(9.1)	32.7	35.0
	<b>Girls</b>	209	33.6	(9.2)	32.3	34.9
BMI (kg/m <sup>2</sup> )	<b>Total</b>	459	17.23	(2.68)	16.99	17.49
	<b>Boys</b>	250	17.39	(2.87)	17.0	17.8
	<b>Girls</b>	209	17.04	(2.43)	16.7	17.4

**Table 2.2** Athletic Skills Tracks per KTK category.

		<b>N</b>	<b>Mean</b>	<b>(± SD)</b>
<b>AST-1 / KTK category</b> (sec)	<b>Total</b>	462	44	(11)
	<b>SMD</b>	6	67	(19)
	<b>MMD</b>	38	52	(14)
	<b>NMG</b>	322	44	(9)
	<b>GMG</b>	89	38	(7)
	<b>HMG</b>	7	35	(5)
<b>AST-2 / KTK category</b> (sec)	<b>Total</b>	463	45	(12)
	<b>SMD</b>	6	79	(16)
	<b>MMD</b>	39	57	(13)
	<b>NMG</b>	322	45	(10)
	<b>GMG</b>	89	38	(7)
	<b>HMG</b>	7	43	(18)

High motor giftedness = HMG, good motor giftedness = GMG, normal motor giftedness = NMG, moderate motor disorder = MMD, serious motor disorder = SMD.

**Table 2.3** Fundamental movement skills per test and age group.

Age (years)		KTK (MQ)			AST-1 (sec)			AST-2 (sec)		
		Total	Boys	Girls	Total	Boys	Girls	Total	Boys	Girls
6	<b>N</b>	32	14	18	32	14	18	32	14	18
	<b>Mean</b>	110	110	111	55	52	57	54	54	54
	<b>(± SD)</b>	(13)	(15)	(11)	(10)	(8)	(11)	(11)	(14)	(9)
7	<b>N</b>	69	41	28	68	40	28	69	41	28
	<b>Mean</b>	105	104	106	51	48	55	50	47	54
	<b>(± SD)</b>	(14)	(14)	(15)	(11)	(7)	(13)	(12)	(10)	(14)
8	<b>N</b>	74	45	29	74	45	29	74	45	29
	<b>Mean</b>	105	103	108	44	43	46	46	45	48
	<b>(± SD)</b>	(12)	(12)	(13)	(8)	(9)	(8)	(9)	(8)	(10)
9	<b>N</b>	87	47	40	87	47	40	87	47	40
	<b>Mean</b>	100	97	103	44	43	46	47	48	46
	<b>(± SD)</b>	(15)	(15)	(14)	(10)	(11)	(8)	(12)	(14)	(10)
10	<b>N</b>	81	40	41	81	40	41	81	40	41
	<b>Mean</b>	103	101	105	41	39	43	41	41	41
	<b>(± SD)</b>	(14)	(15)	(12)	(11)	(8)	(13)	(10)	(10)	(9)
11	<b>N</b>	86	45	41	86	45	41	86	45	41
	<b>Mean</b>	105	102	108	37	37	38	39	39	39
	<b>(± SD)</b>	(13)	(12)	(13)	(6)	(6)	(6)	(8)	(9)	(7)
12	<b>N</b>	34	20	14	34	20	14	34	20	14
	<b>Mean</b>	102	103	101	37	35	40	43	41	45
	<b>(± SD)</b>	(15)	(17)	(12)	(9)	(9)	(7)	(15)	(16)	(13)
<b>Total</b>	<b>N</b>	463	252	211	462	251	211	463	252	211
	<b>Mean</b>	104	102	106	44	42	46	45	44	46
	<b>(± SD)</b>	(14)	(14)	(13)	(11)	(10)	(12)	(12)	(12)	(11)

KTK = Körperkoordinatinon Test für Kinder Motor Quotient, AST-1 = athletic skills track 1, AST-2 = athletic skills track.

## 2.4.1 Learning effect ASTs

Both AST were measured twice to examine if there was a learning effect. The intraclass correlation coefficient between the first and second trial was 0.875 (95% CI [0.852–0.895]) for AST-1 and 0.891 (95% CI [0.870–0.908]) for AST-2 indicating a high degree of test–retest reliability (Shrout & Fleiss, 1979). However, a paired sample t-test showed a small, but significant learning effect between the two trials. On average, children completed the second trial 2 s faster than the first trial (AST-1 trial 1:  $45 \pm 11$  s, versus trial 2:  $44 \pm 11$  s;  $t = 6.026$ ,  $P < 0.05$ ; AST-2 trial 1:  $47 \pm 12$  s versus trial 2:  $45 \pm 12$  s;  $t = 8.226$ ,  $P < 0.05$ ). Because of the small difference between the two trials it was decided to continue with the data of the second trial of AST-1 and AST-2 to examine their relationship with the KTK.

## 2.4.2 Concurrent validity ASTs

Overall, there was a low correlation between AST-1 and the KTK ( $r = -0.474$ ,  $P < 0.01$ ), and a moderate correlation between AST-2 and the KTK ( $r = -0.502$ ,  $P < 0.01$ ) (see Table 2.4). The correlations between AST-1 and KTK were higher when split up for gender (girls:  $r = -0.501$ ,  $P < 0.01$ ; boys:  $r = -0.533$ ,  $P < 0.01$ ). For AST-2 the correlation between AST-2 and the KTK was lower for girls ( $r = -0.448$ ,  $P < 0.01$ ) and higher for boys ( $r = -0.566$ ,  $P < 0.01$ ) than the overall correlation. In general, correlations were also higher when examined per age group (see also Figure 2.3 and 2.4). With the exception of the low correlation coefficient of the AST-2 in 7-year-olds ( $r = -0.290$ ,  $P < 0.01$ ), the other correlation coefficients were near or far above 0.50 for each age group.

**Table 2.4** Pearson correlations between KTK and AST-1, AST-2 per gender.

		AST-1 (sec)		AST-2 (sec)			
		Total	Boys	Girls	Total	Boys	Girls
KTK	Total	1	-0.474**		-0.502**		
(MQ)	Boys		-0.533**			-0.566**	
	Girls			-0.501**			-0.448**

\*\* Correlation is significant at the 0.01 level (2-tailed). KTK = Körperkoordinations Test für Kinder, MQ = Motor Quotient, AST-1 = athletic skills track 1, AST-2 = athletic skills track 2.



### 2.4.3 Discriminative ability ASTs

A one-way ANOVA revealed that both ASTs were able to categorise children's FMS into high, good or normal motor giftedness or moderate or serious motor disorder as indicated by the KTK. The time scores on AST-1 and AST-2 were significantly different between the five KTK categories (AST-1: Welch's  $F(4,21.011) = 22.968, P < 0.05$ ; AST-2: Welch's  $F(4,20.366) = 27.746, P < 0.05$ ). On average, children with a higher motor giftedness completed the ASTs significantly faster than children with lower levels of FMS (see also Table 2.2). Tukey HSD post hoc analysis revealed that the differences between the KTK categories were all significant ( $P < 0.05$ ) for AST-1 and AST-2 except for the difference between NMG–HMG (AST-1: mean difference NMG–HMG: 9.252 with 95% CI  $-0.107$ – $19.57$ ; AST-2: mean difference NMG–HMG: 1.392 with 95% CI  $-9.62$ – $12.40$ ) and GMG–HMG (AST-1: mean difference GMG–HMG:  $-2.693$  with 95% CI  $-7.91$ – $13.30$ ; AST-2: mean difference GMG–HMG:  $-5.19$  with 95% CI  $-16.50$ – $6.12$ ).

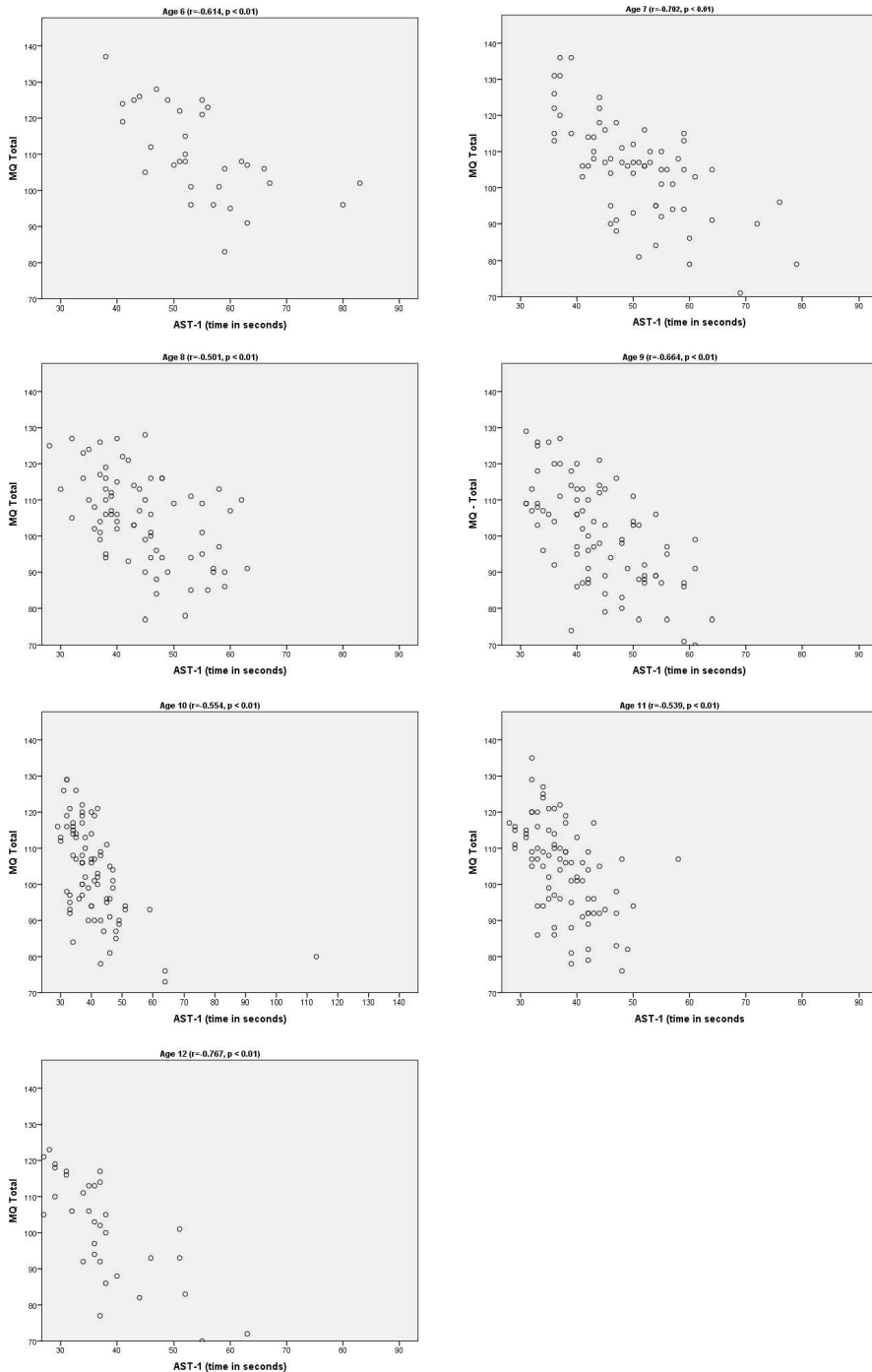


Figure 2.3 Scatter plots and Pearson correlations between MQ KTK and AST-1 per age group.

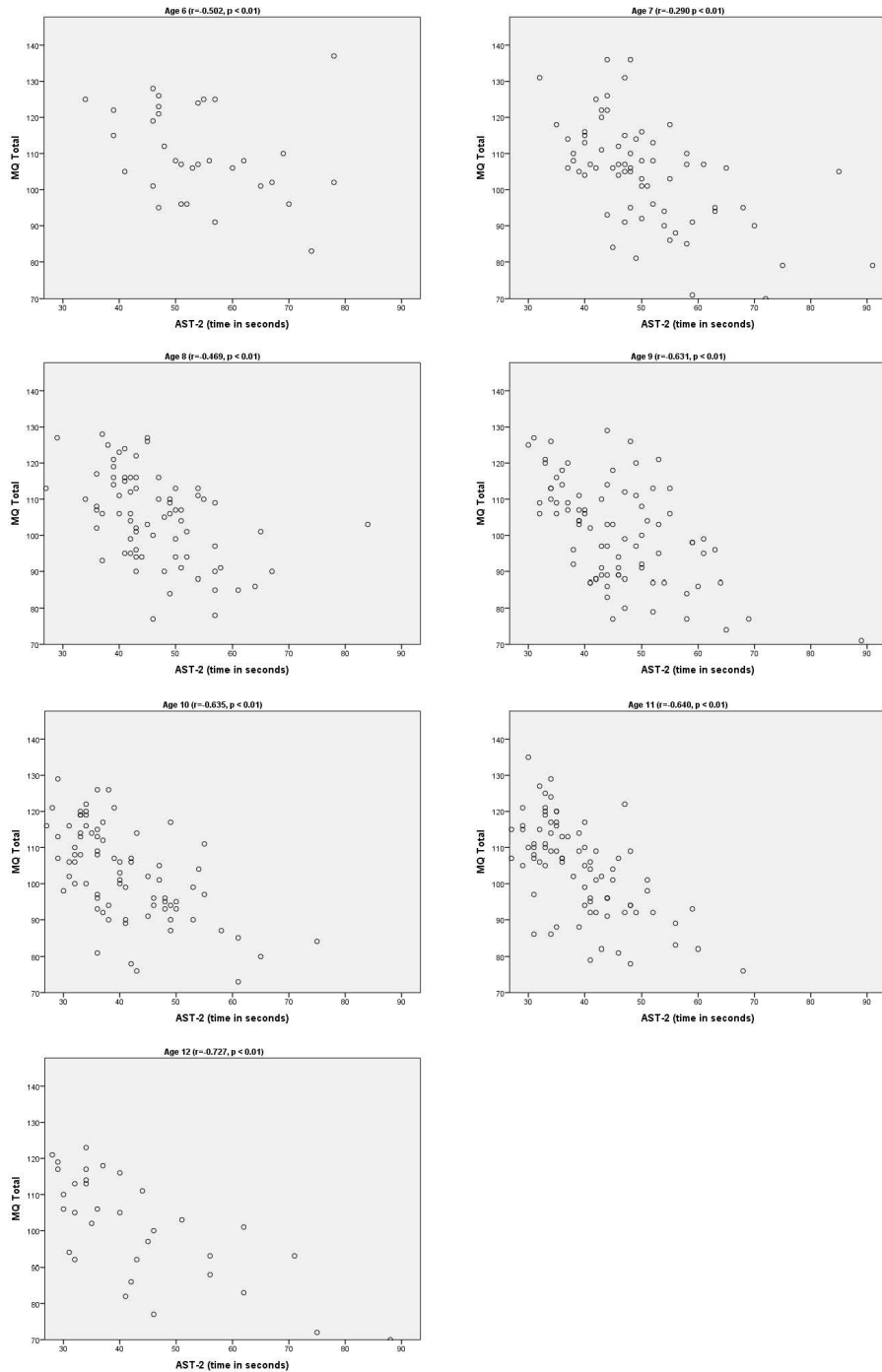


Figure 2.4 Scatter plots and Pearson correlations between MQ KTK and AST-2 per age group.

## 2.5 Discussion

In the last decade there is an increasing interest in children's physical activity level at a young age and its relationship with an active lifestyle later on. Therefore, measuring physical activity, as well as children's fitness level and motor competence level is now prominently featured within the curriculum of PE (Cale et al., 2014). In this study, a novel screening tool – the AST – has been refined and evaluated in order to provide PE teachers a valid, feasible and fun way to determine children's motor competence without them being aware that they are “evaluated”. The results show that an AST is a feasible screening tool in a PE setting. The track measures a broad range of FMS using sports equipment and materials that are available in every gym in The Netherlands. On average, it took less than 1 min per child to complete one of the tests (AST-1 or AST-2). A complete class (25–30 children) can thus be measured in one regular PE lesson of 50 min. Although a learning effect was found between the first and second trials (after three try-out trials), the test–retest reliability of both AST-1 and AST-2 was high. Furthermore, both AST-1 and AST-2 seem to be able to measure FMS in children aged 6–12 years as the correlation between the time to complete the tracks and the KTK MQ was moderate to high in most age groups.

In addition, the results show that the tracks are able to distinguish on group level between serious motor disorder, moderate motor disorder and normally to highly motor gifted children. When comparing our results with those presented in the review of Cools et al. (2009), it can be concluded that the validity of both tests (AST-1 and AST-2) is in the same range as the most frequently used assessment tools (range  $r$ : 0.43–0.87). However, the AST is more feasible in a PE setting than the MOT 4–6, Movement-ABC, PDMS, BOTMP, TGMD, MMT and the KTK. In other words, the AST seems to be a valuable alternative for the current motor competence tests to assess children's FMS in a PE setting. Another advantage of the AST is that it can be implemented within the context of the regular PE lessons (Cale et al., 2014).

In order to increase its value, it is recommended to develop age-specific ASTs in future studies since the current tracks were less distinct for the more gifted children. This may in part be due to the small sample sizes in the good to high motor gifted groups. However, a trend towards a difference between these groups was apparent, which may be indicative for a more pronounced difference in assessments with a larger sample size. Next, certain constraints cannot be ruled out at this point. Further study is required to establish the discriminative ability of the AST for more gifted children. The AST could be adjusted by using different tracks for different age groups (for instance by taking into account coordinative abilities per age group), providing more information for the younger age groups and increasing the discriminative ability of the

test. If, after further refinement, the AST would be capable of distinguishing between different motor skill categories, this track could also be used to identify potential athletes in different sports. Vandorpe et al. (2012) concluded that a non-sports-specific motor competence test could contribute positively in the process of talent identification. In addition, a motor competence test could also provide insight into talent development (Faber, Nijhuis-Van Der Sanden, Elferink-Gemser, & Oosterveld, 2014). PE teachers and physiotherapists might also use the AST to identify the less motor gifted children. It is estimated that 5–10% of children are diagnosed with a Developmental Coordination Disorder (Kirby & Sugden, 2007). Based on the differences between the seriously motor disordered, the moderately motor disordered and the normally motor gifted group, the AST might be used as a screening tool to identify those children at risk for motor competence disorder such as DCD. Subsequently, AST screening should be followed by a more comprehensive diagnostic assessment. Since FMS are founded in early childhood, and FMS are predictors of people's physical activity level later in life (Lloyd et al., 2014), it is important to identify children with low motor coordination as early as possible (Lopes et al., 2012; Runhaar et al., 2010). By using the AST as a screening tool in PE, it is possible to identify this group in early childhood.

This study had several limitations. First, in this study, only time to complete the track was measured as an indicator of a child's motor skill level. Although this is a very feasible measure in a PE setting, there is no knowledge about the relation between time to complete a series of fundamental movement skills and the quality of the performed fundamental movement skills within the track. This may help to further refine the tracks. Second, no information was gathered on children's sports participation or physical activity level. Children who perform certain sports, such as soccer or basketball, might have been at an advantage since AST-1 consisted partly of manipulative skills, thereby influencing the correlation between the track and the KTK. In future studies, preferably, objective assessment methods should be employed to measure children's physical activity level as well as their motor skill level. The influence of children's BMI on the correlation between AST and the KTK also needs to be studied, since D'Hondt et al. (2011) showed that childhood overweight results in poorer KTK performances. Third, this study is cross-sectional. The responsiveness of the AST to change in the course of time needs to be established. This could be a focus in future studies. In addition, limited aspects of reliability and validity have been subject of inquiry in the present study. Other features of validity (such as ecological validity) and reliability (inter-rater reliability), and also the within person error and minimal detectable change, need to be addressed. In case the AST holds true

with a reference to identify deviations from the norm.

### 2.5.1 Conclusion

In conclusion, this study shows that an AST might be a feasible alternative for the current motor competence tests to assess children's FMS in a PE setting. The tracks showed moderate to high correlation in most age groups compared to a frequently used but time-consuming motor competence test and provided an indication of children's motor skills level. Future studies should further assess the reliability, discriminative ability and validity of age-specific versions of the AST.

## 2.6 References

Bouffard, M., Watkinson, J. E., Thompson, L. P., Causgrove Dunn, J. L., & Romanow, S. K. E. (1996). A test of the activity deficit hypothesis with children with movement difficulties. *Adapted Physical Activity Quarterly*, 13, 61–73.

Cale, L., Harris, J., & Chen, M. H. (2014). Monitoring health, activity and fitness in physical education: Its current and future state of health. *Sport, Education and Society*, 19(4), 376–397. doi:10.1080/13573322.2012.681298

Cantell, M., Crawford, S. G., & Doyle-Baker, P. K. (2008). Physical fitness and health indices in children, adolescents and adults with high or low motor competence. *Human Movement Science*, 27(2), 344–362. doi:10.1016/j.humov.2008.02.007

Centers for Disease Control and Prevention. (2011). School health guide- lines to promote healthy eating and physical activity. *MMWR. Recommendations and Reports: Morbidity and Mortality Weekly Report. Recommendations and Reports/Centers for Disease Control*, 60(RR–5), 1.

Clark, J. E., & Metcalfe, J. S. (2002). The mountain of motor development: A metaphor. *Motor Development: Research and Reviews*, 2, 163–190.

Cliff, D. P., Okely, A. D., Smith, L. M., & McKeen, K. (2009). Relationships between fundamental movement skills and objectively measured physical activity in preschool children. *Pediatric Exercise Science*, 21(4), 436–449.

Cools, W., Martelaer, K. D., Samaey, C., & Andries, C. (2009). Movement skill assessment of typically developing preschool children: A review of seven movement skill assessment tools. *Journal of Sports Science & Medicine*, 8(2), 154–168.

D'Hondt, V., Deforche, B., Vaeyens, R., Vandorpe, B., Vandendriessche, J., Pion, J., & Lenoir, M. (2011). Gross motor coordination in relation to weight status and age in 5- to 12-year-old boys and girls: A cross- sectional study. *International Journal of Pediatric Obesity*, 6(sup3), e556– e564.

Dollman, J., Norton, K., & Norton, L. (2005). Evidence for secular trends in children's physical activity behaviour. *British Journal of Sports Medicine*, 39(12), 892–897; discussion 897. doi:10.1136/bjism.2004.016675

Faber, I. R., Nijhuis-Van Der Sanden, M. W., Elferink-Gemser, M. T., & Oosterveld, F. G. (2015). The Dutch motor skills assessment as tool for talent development in table tennis: A reproducibility and validity study. *Journal of Sports Sciences*, 33(11), 1149–1158.

Haga, M. (2008). The relationship between physical fitness and motor competence in children. *Child: Care, Health and Development*, 34(3), 329–334. doi:10.1111/j.1365-2214.2008.00814.x

Hallal, P. C., Andersen, L. B., Bull, F. C., Guthold, R., Haskell, W., & Ekelund, U. (2012). Global physical activity levels: Surveillance progress, pitfalls, and prospects. *The Lancet*, 380(9838), 247–257. doi:10.1016/S0140-6736(12)60646-1

Hardy, L., King, L., Espinel, P., Cosgrove, C., & Bauman, A. (2013). NSW schools physical activity and nutrition survey 2010: Full report. Sydney: SPANS.

Hildebrandt, V. H., Bernaards, C. M., & Stubbe, J. H. (2013). *Trendrapport bewegen en gezondheid, 2010/2011*. Leiden: TNO.

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

Kiphard, E. J., & Schilling, F. (2007). *Körperkoordinationstest für Kinder. 2. Überarbeitete und ergänzte Auflage*. Weinheim: Beltz Test GmbH.

Kirby, A., & Sugden, D. A. (2007). Children with developmental coordination disorders. *Journal of the Royal Society of Medicine*, 100(4), 182–186. doi:10.1258/jrsm.100.4.182

Lenoir, M., Vanderpe, B., D'Hondt, E., Pion, J., Vandendriessche, J., Vaeyens, R., & Philippaerts, R. M. (2014). *KTK-NL Körperkoordinationstest für Kinder. Herwerkte, gehernormeerde en vertaalde uitgave van de KTK voor het Nederlandstalig gebied (Vol. 1, pp. 80)*. Destelbergen: Universiteit Gent.



Lloyd, M., Saunders, T. J., Bremer, E., & Tremblay, M. S. (2014). Long-term importance of fundamental motor skills: A 20-year follow-up study. *Adapted Physical Activity Quarterly*, 31(1), 67–78. doi:10.1123/apaq.2013-0048

Lopes, V. P., Stodden, D. F., Bianchi, M. M., Maia, J. A., & Rodrigues, L. P. (2012). Correlation between BMI and motor coordination in children. *Journal of Science and Medicine in Sport*, 15(1), 38–43. doi:10.1016/j.jsams.2011.07.005

Lubans, D. R., Morgan, P. J., Cliff, D. P., Barnett, L. M., & Okely, A. D. (2010). Fundamental movement skills in children and adolescents: Review of associated health benefits. *Sports Medicine*, 40(12), 1019–1035. doi:10.2165/11536850-000000000-00000

Okely, A. D., Booth, M. L., & Patterson, J. W. (2001). Relationship of physical activity to fundamental movement skills among adolescents. *Medicine and Science in Sports and Exercise*, 33(11), 1899–1904. doi:10.1097/00005768-200111000-00015

Reijgersberg, N., van der Werff, H., & Lucassen, J. (2013). Nulmeting bewegingsonderwijs: Onderzoek naar de organisatie van het bewegingsonderwijs in het primair onderwijs. Utrecht: Mulier Instituut.

Runhaar, J., Collard, D. C., Singh, A. S., Kemper, H. C., Van Mechelen, W., & Chinapaw, M. (2010). Motor fitness in Dutch youth: Differences over a 26-year period (1980-2006). *Journal of Science and Medicine in Sport*, 13 (3), 323–328. doi:10.1016/j.jsams.2009.04.006

Salmon, J., & Timperio, A. (2007). Prevalence, trends and environmental influences on child and youth physical activity. *Pediatric Fitness*, 50, 183–199.

Shrout, P. E., & Fleiss, J. L. (1979). Intraclass correlations: Uses in assessing rater reliability. *Psychological Bulletin*, 86(2), 420–428. doi:10.1037/0033-2909.86.2.420

Stodden, D. F., Goodway, J. D., Langendorfer, S. J., Roberton, M. A., Rudisill, M. E., Garcia, C., & Garcia, L. E. (2008). A developmental perspective on the role of motor skill competence in physical activity: An emergent relationship. *Quest*, 60(2), 290–306. doi:10.1080/00336297.2008.10483582

Stodden, D. F., Langendorfer, S. J., & Robertson, M. A. (2009). The association between motor skill competence and physical fitness in young adults. *Research Quarterly for Exercise and Sport*, 80(2), 223–229. doi:10.1080/02701367.2009.10599556

Vandorpe, B., Vandendriessche, J., Lefèvre, J., Pion, J., Vaeyens, R., Matthys, S., Lenoir, M. (2011). The KörperkoordinationsTest für Kinder: Reference values and suitability for 6–12-year-old children in Flanders. *Scandinavian Journal of Medicine & Science in Sports*, 21(3), 378–388. doi:10.1111/sms.2011.21.issue-3

Vandorpe, B., Vandendriessche, J., Vaeyens, R., Pion, J., Lefèvre, J., Philippaerts, R. M., & Lenoir, M. (2012). The value of a non-sport-specific motor test battery in predicting performance in young female gymnasts. *Journal of Sports Sciences*, 30(5), 497–505. doi:10.1080/02640414.2012.654399

Wormhoudt, R., Teunissen, J. W., & Savelsbergh, G. J. P. (2012). *Athletic skills model*. Nieuwegein: Arko Sports Media.

Wrotniak, B. H., Epstein, L. H., Dorn, J. M., Jones, K. E., & Kondilis, V. A. (2006). The relationship between motor proficiency and physical activity in children. *Pediatrics*, 118(6), e1758–e1765. doi:10.1542/peds.2006-0742





# 3

## Reliability and concurrent validity of a motor skill competence test among 4- to 12-year old children

Hoeboer, J., Krijger-Hombergen, M., Savelsbergh, G., & De Vries, S. (2018).

*Journal of Sports Sciences, 36(14), 1607-1613.*

The purpose of this study was to examine the test-retest reliability, internal consistency and concurrent validity of the Athletic Skills Track (AST). During a regular PE lesson, 930 4- to 12-year old children (448 girls, 482 boys) completed two motor skill competence tests: (1) the Körperkoordination-Test für Kinder (KTK) and (2) an age-related version of the AST (age 4–6 years: AST-1, age 6–9 years: AST-2, and age 9–12 years: AST-3). The test-retest reliability of the AST was high (AST-1: ICC = 0.881 (95% CI: 0.780–0.934); AST-2: ICC = 0.802 (95% CI: 0.717–0.858); and AST-3: ICC = 0.800 (95% CI: 0.669–0.871)). The internal consistency, concerning the three age-bands of the AST was above the acceptable level of Cronbach's  $\alpha > 0.70$  (AST-1:  $\alpha = 0.764$ ; AST-2:  $\alpha = 0.700$ ; and AST-3:  $\alpha = 0.763$ ). There was a moderate to high correlation between the time to complete the AST, and the age- and gender-related motor quotients of the KTK (AST-1:  $r = -0.747$ ,  $P = 0.01$ ; AST-2:  $r = -0.646$ ,  $P = 0.01$ ; and AST-3:  $r = -0.602$ ,  $P = 0.01$ ). The Athletic Skills Track is a reliable and valid assessment tool to assess motor skill competence among 4- to 12-year old children in the PE setting.

### 3 Reliability and concurrent validity of a motor skill competence test among 4- to 12-year old children

#### 3.1 Introduction

Being motor competent allows children to be active in a diversity of physical activities (Clark & Metcalfe, 2002; Seefeldt, 1980). Motor competence is a global term used to describe goal-directed human movement (Robinson et al., 2015). According to Clark and Metcalfe (2002) and Stodden et al. (2008) the development of fundamental movement skills (FMS) in early childhood is the precursor of motor skill competence. Different categories of FMS are: locomotive skills (e.g., hopping and sliding), manipulative skills (e.g., throwing and kicking), also referred to as object control skills, and stability skills (e.g., balance and body rolling) (Gallahue, Ozmun, & Goodway, 2012). Several studies have shown that children's motor skill competence is positively associated with their physical activity level (Lubans, Morgan, Cliff, Barnett, & Okely, 2010; Robinson et al., 2015). Motor skill competence is also positively related to people's physical activity level in adolescence (Barnett, Van Beurden, Morgan, Brooks, & Beard, 2009; Lima et al., 2017). In a longitudinal study, a relationship has been found between motor skill competence at age 6 and physical activity at age 26 (Lloyd, Saunders, Bremer, & Tremblay, 2014). A high level of motor skill competence in youth is also associated with several health-related parameters such as lower body mass index (D'Hondt et al., 2011) and better cardiorespiratory fitness (Okely, Booth, & Patterson, 2001).

Even though an active lifestyle is related to well-known health effects (Janssen & LeBlanc, 2010; Pate et al., 1995), physical activity levels among children and adolescents seem to be decreasing in many countries (Hallal et al., 2012; Tremblay et al., 2014). In the Netherlands physical activity levels are also decreasing (Van Mechelen, Twisk, Post, Snel, & Kemper, 2000). Motor skills levels appear to have dropped as well over the last decades among Dutch youth (Runhaar et al., 2010). To increase our understanding of the development of motor skill competence during childhood and factors that can influence the development, a reliable and valid assessment tool is needed. In recent decades, several motor skill competence tests have been developed. Some tests have a quantitative outcome and are product orientated (i.e., number of jumps in a specific time) such as the Körperkoordination-Test für Kinder (KTK) (Kiphard & Schilling, 2007). Other tests are process orientated and have a qualitative outcome (i.e., demonstrating specific components while jumping) such as the Movement Assessment Battery for Children-2 (Movement-ABC-2) (Henderson, Sugden, & Barnett, 2007), and the

Test of Gross Motor Development (TGMD) (Ulrich, 2000). Empirical evidence shows a moderate to strong association between process and product orientated assessments (Logan, Barnett, Goodway, & Stodden, 2017; Miller, Vine, & Larkin, 2007). Most of the motor skill competence tests have been developed for clinical purposes and are used to identify children with motor impairment or medical deficits, such as developmental coordination disorder (DCD) (Lenoir et al., 2014; Tidén, Lundqvist, & Nyberg, 2015). These clinically based tests are not very feasible in a physical education (PE) setting (Cools, De Martelaer, Samaey, & Andries, 2009). Most clinically based motor skill competence tests are very time-consuming since it takes at least 20 minutes to measure one child (Cools et al., 2009). This makes it difficult to integrate the measurement of all children in one class in a single, one-hour, PE lesson. In addition, special test materials are needed for some tests, for example; to execute the KTK you will need specialized material, such as three balance beams with different widths. Despite the fact that PE teachers would like to monitor the motor skill competence of children more objectively (Lander, Morgan, Salmon, & Barnett, 2016), the mentioned shortcomings seem to be the reason many PE teachers currently do not use motor skill competence tests structurally.

When selecting an assessment tool, the possibility of negative experiences of children being tested also needs to be taken into account. A negative test experience can have a negative impact on the motivation of the children with a negative experience, especially those most in need of encouragement (Cale, Harris, & Chen, 2014). Cale et al. (2014) formulated a number of recommendations to be taken into account when testing children in the PE setting, in order to reduce this potential negative impact of testing. They recommend testing within the context of a regular PE lesson instead of in an externally arranged test lesson. They also state that it is important that the test is a positive experience for all children (Cale et al., 2014).

To address those shortcomings, a novel, quick, convenient, and low-cost motor skill competence test, the Athletic Skills Track (AST), has been developed to assess motor skill competence among 4- to 12-year-old children in a PE setting (Hoeboer et al., 2016). The AST makes it possible to assess 25–30 children in a one-hour PE lesson. The AST is fundamentally different than most conventional motor skill competence tests. Rather than assessing isolated movements, the track consists of a series of 5–7 concatenated FMS (e.g., hopping and balancing), providing a motor skill assessment of children's motor skill competence. This idea is based on the concept that motor coordination (e.g., coupling, spatial orientation, and balance ability) is crucial in developing motor competence (Wormhoudt, Teunissen, & Savelsbergh,

2012). Motor coordination can be defined as the ability to have body segments work together in an organized manner (Turvey, 1990) and are considered an underlying component of FMS (Barnett et al., 2016).

The track is constructed using equipment available in every gymnasium in The Netherlands. Children are asked to complete the track as fast as possible. The AST has a quantitative measurement outcome, i.e., the time a child needs to complete the track. A previous study among 463 6- to 12-year old children has shown that the AST is a valid motor skill competence test in a PE setting (Hoeboer et al., 2016). In this study the time to complete the AST was correlated to the Motor Quotient (MQ) of the KTK. When split up by age group, the associations ranged between  $r = -0.502$  ( $P < 0.01$ ) and  $r = -0.767$  ( $P < 0.01$ ). So, the faster a child completed the AST, the higher their MQ on the KTK. However, a ceiling effect was found among the oldest children. Therefore, the purpose of the present study was to examine the reliability (test-retest reliability and internal consistency) and concurrent validity of an age-related version of the AST (age 4–6 years: AST-1, age 6–9 years: AST-2, and age 9–12 years: AST-3) to assess motor skill competence of children aged 4–12 years in a PE setting.

## 3.2 Methods

This study consisted of two sub-studies. In the first study (study 1), the test-retest reliability of the three age-bands of the AST was examined. In the second study (study 2), the internal consistency and concurrent validity of the age-related AST were investigated. Permission for this study was granted by the Ethical Committee of the Faculty of Human Movement Sciences, VU University Amsterdam, The Netherlands (ECB 2015-31).

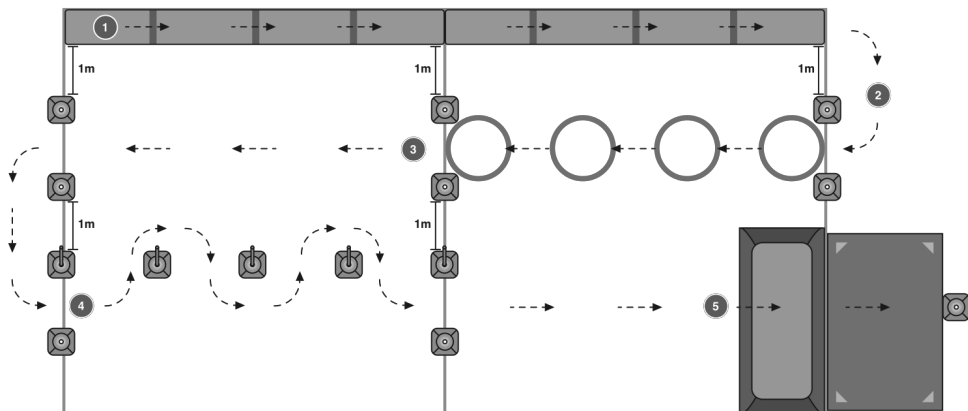
### 3.2.1 Participants

Fourteen primary schools spread over the The Hague region took part in this study. The random selected schools are internship schools of the The Hague University of Applied Sciences (THUAS). For study 1, ten PE teachers from ten different elementary schools were recruited. For study 2, four other elementary schools were recruited. Written informed consent was obtained from the parents of the children (grade 1–8). All parents were informed about the purpose and nature of the study, and that their children could withdraw from the study at any time.

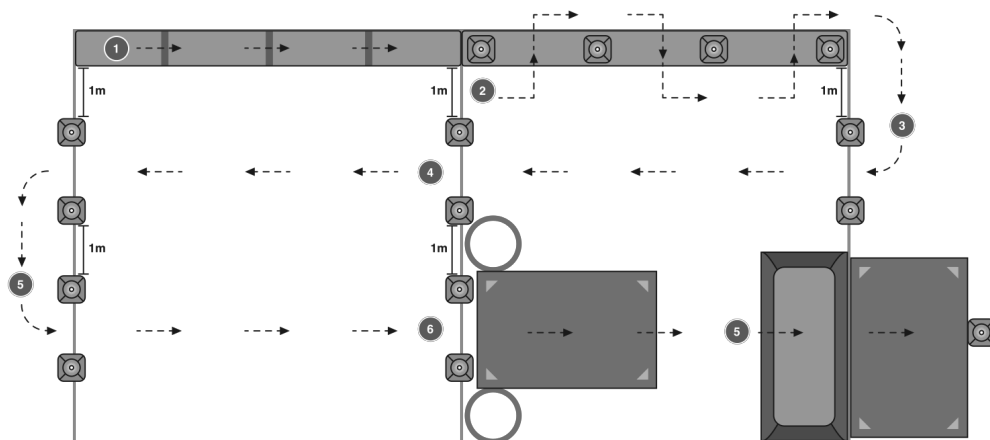


### 3.2.2 Procedure

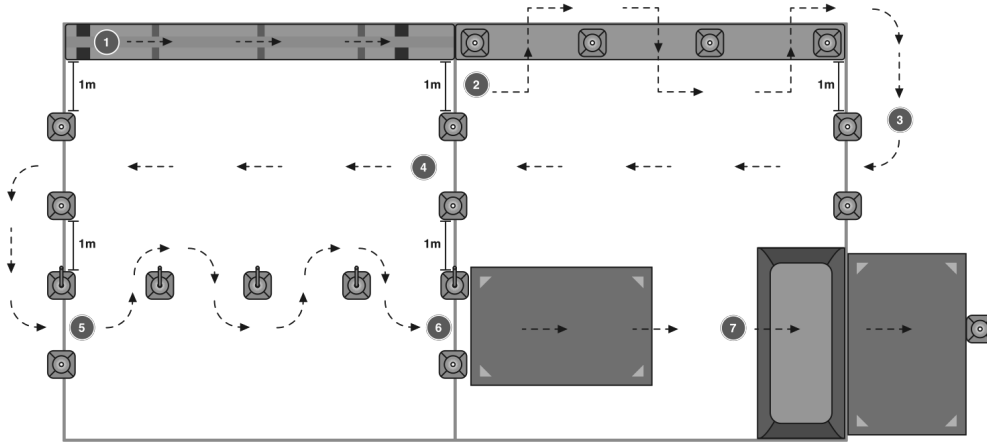
All measurements were executed between February and March 2015. In both studies children performed the AST. Three age-related tracks (AST-1, AST-2 and AST-3) were developed with a team of 10 PE teachers. In 4 sessions of 2 hours the tracks were designed based on the experience of PE teachers and the former study on the AST (Hoeboer et al., 2016). The validation study of Hoeboer et al. (2016) laid the foundations of the new tracks. The tracks include a string of different FMS (AST-1:  $n = 5$ , AST-2:  $n = 7$ , and AST-3:  $n = 7$ ) to be completed as quick as possible. In all three tracks the same FMS are tested, with the difficulty of the tasks ascending from AST-1 to AST-3 (see Figure 3.1, 3.2 and 3.3). AST-1 has been designed for the youngest children in the age of 4–6 years and consists of the following skills: (1) Walking, (2) Traveling jumps, (3) Alligator crawl, (4) Slaloming, and (5) Clambering. Children between 6 and 9 years performed AST-2. In this track the following skills had to be performed: (1) Walking, (2) Traveling jumps, (3) Hopscotch, (4) Alligator crawl (backwards), (5) Running (backwards), (6) Pencil roll, and (7) Clambering. Children in the age of 9–12 years completed AST-3 consisting of: (1) Walking (backwards), (2) Traveling jumps, (3) Hopscotch (backwards), (4) Alligator crawl (backwards), (5) Slaloming (backwards), (6) Forward roll, and (7) Clambering.



**Figure 3.1** AST-1 schematically displayed. (1) Walking, (2) Traveling jumps, (3) Alligator crawl, (4) Slaloming, and (5) Clambering



**Figure 3.2** AST-2 schematically displayed. (1) Walking, (2) Traveling jumps, (3) Hopscotch, (4) Alligator crawl (backwards), (5) Running (backwards), (6) Pencil roll, and (7) Clambering.



**Figure 3.3** AST-3 schematically displayed. (1) Walking (backwards), (2) Traveling jumps, (3) Hopscotch (backwards), (4) Alligator crawl (backwards), (5) Slaloming (backwards), (6) Forward roll, and (7) Clambering.

The age-related tracks were designed because the former study, with one track for all age groups, showed a ceiling effect in the older age groups (Hoeboer et al., 2016). In both studies, the PE teacher (study 1) and research assistant (study 2) showed how to complete the track. After the example of the teacher or assistant the children completed three try-out trials. During the try-out trials the PE teacher (study 1) or the research assistant (study 2) gave feedback to the children about how the skills should be performed so that the children were ready to execute the measurement trial independently. The children performed one (study 1,  $n = 717$ ) or two (study 2,  $n = 213$ ) measurement trials on each measurement day. More information about the procedure of the AST can be found in the test manual which is presented in appendix 2.

### 3.2.3 Measurements

To examine the test-retest reliability (study 1) of the AST, all children performed the AST twice between February and March 2015 two weeks apart. The PE teacher conducted the measurements during a regular PE lesson in one-third of the PE classroom. In the other two-thirds, the PE lesson was conducted as planned by the PE teacher. The PE teacher used a stopwatch to assess the time children needed to complete the track. All PE teachers who participated as test leaders in study 1 were trained in conducting the test according to the AST protocol (see appendix 2) by the corresponding author during a two-hour training session where they practiced the performance of the track and the measurement.

To examine the internal consistency of the AST (study 2) a Freelap Timing® system was used. With an accuracy of 2/100", this system was used to measure split times of different movement skill tasks in the track automatically in addition to the total time to complete the track. Besides the AST, children also performed the KTK in study 2 to examine the concurrent validity of the AST. The origin of the KTK can be found in Germany (Kiphard & Schilling, 2007; Vandorpe et al., 2011). For the KTK, the test protocol for the Dutch language area (Lenoir et al., 2014) was followed. The KTK contains four subtests, i.e., moving across the floor in 20 seconds by stepping from one plate to the next, jumping from one leg over an increasing pile of pillows, jumping laterally as many times as possible over a wooden slat and walking backwards on a balance beam. Both the AST and the KTK were conducted in a separate section of the gym by a team of four research assistants (fourth year PE students of THUAS). All research assistants were trained in conducting the KTK and the AST according to the protocols by the corresponding author in four meetings of approximately 45 minutes. Testing was spread out over a two-week time period in February 2015. To measure all the children in this period a uniform schedule was developed: during the first week, the children conducted the AST, based on their age. Next, the children completed the KTK within this two-week time period.

#### 3.2.4 Data analysis

Of the 1284 children who participated in the studies, 930 children met the inclusion criteria (i.e., age: between 4 and 12 years; all tests completed). The data were at first registered in Excel (Microsoft Office professional plus, 2013) and then analysed using the SPSS software (IBM SPSS 22.0 64-bit edition). To check for normality of the outcome parameters in the sample (i.e., time and split-time to complete the AST), histograms were plotted and the kurtosis and skewness values for each of the outcome parameters were assessed.

An Independent Samples T-test was used to investigate possible differences between boys and girls. Values were considered statistically significant at  $P < 0.05$ . To examine the test – retest reliability of the AST the intraclass correlation coefficient (ICC) was calculated (Altman, 1991). ICC values below 0.20 reflect poor agreement, between 0.21 and 0.40 fair agreement, between 0.41 and 0.60 moderate agreement, between 0.61 and 0.80 good agreement, and higher than 0.80 very good agreement (Altman, 1991). In addition, a “quality control” of the level of agreement between the two trials of the AST was performed by constructing Bland and Altman plots (Bland & Altman, 1986). In these plots the difference between the two trials is plotted against the mean of the two trials. The limits of agreement (LoA), proposed as a parameter of

measurement error by Bland and Altman (1986), were represented in the plots as well as the change with increasing mean values (Bland & Altman, 1999).

The internal consistency of the AST was examined by computing Cronbach's  $\alpha$  and the corrected item-total correlation regarding the split time to complete the five to seven different tasks within the AST. The values of Cronbach's  $\alpha$  can be interpreted as follows:  $<0.50$  as unacceptable,  $0.50-0.60$  as poor,  $0.60-0.70$  as questionable,  $0.70-0.80$  as acceptable,  $0.80-0.90$  as good, and  $>0.90$  as excellent (Cronbach, 1951). Since a very small amount of time was required to complete task six and seven within AST-2 and AST-3 the split times of the sixth and seventh task have been added together. It is recommended that the item-total correlation is above 0.20, indicating that the level of item redundancy is acceptable and that the same construct is measured by the other items included (Everitt & Skrondal, 2010; Field, 2005).

The concurrent validity of the AST was examined by calculating Pearson's correlation coefficients between the time to complete the AST and the age- and gender-related motor quotients (MQ) of the KTK (Vandorpe et al., 2011). The values of the correlation coefficient ( $r$ ) can be interpreted as follows: negligible:  $<0.30$ ; low:  $0.30-0.50$ ; moderate:  $0.50-0.70$ ; high:  $0.70-0.90$ ; and very high:  $0.90-1.00$  (Hinkle, Wiersma, & Jurs, 2003).

### 3.3 Results

In total, 717 children (344 girls and 373 boys) participated in study 1 and 213 other children (104 girls and 109 boys) in study 2. Table 3.1 shows the number of participating children and the mean age of these children for each study per track.

**Table 3.1** Age and gender of participants per Athletic Skills Track for each study.

		Study 1			Study 2			
		N	Mean	(SD)	N	Mean	(SD)	
AST-1	Age (years)	Boys	27	6	(1)	23	5	(1)
		Girls	26	6	(1)	16	5	(1)
		Total	53	6	(1)	39	5	(1)
AST-2	Age (years)	Boys	171	8	(1)	34	8	(1)
		Girls	159	8	(1)	38	7	(1)
		Total	330	8	(1)	72	8	(1)
AST-3	Age (years)	Boys	175	11	(1)	52	10	(1)
		Girls	159	11	(1)	50	10	(1)
		Total	334	11	(1)	102	10	(1)
<b>Overall</b>			717	9	(2)	213	9	(2)

AST-1= Athletic Skills Track 1; AST-2= Athletic Skills Track 2; AST-3= Athletic Skills Track 3.

As shown in Table 3.2, in both studies the children completed the AST within 1 minute, independent of the version of the AST. In study 1, the time to complete the track ranged between 14.0 and 35.0 seconds for AST-1, 17.2–57.3 seconds for AST-2 and 17.2–54.4 seconds for AST-3. In study 2, the time to complete the track ranged between 14.4 and 35.0 seconds for AST-1, 20.1–49.8 seconds for AST-2 and 17.0–42.3 seconds for AST-3. In both studies, there were no significant differences found in time to complete the AST between boys and girls.

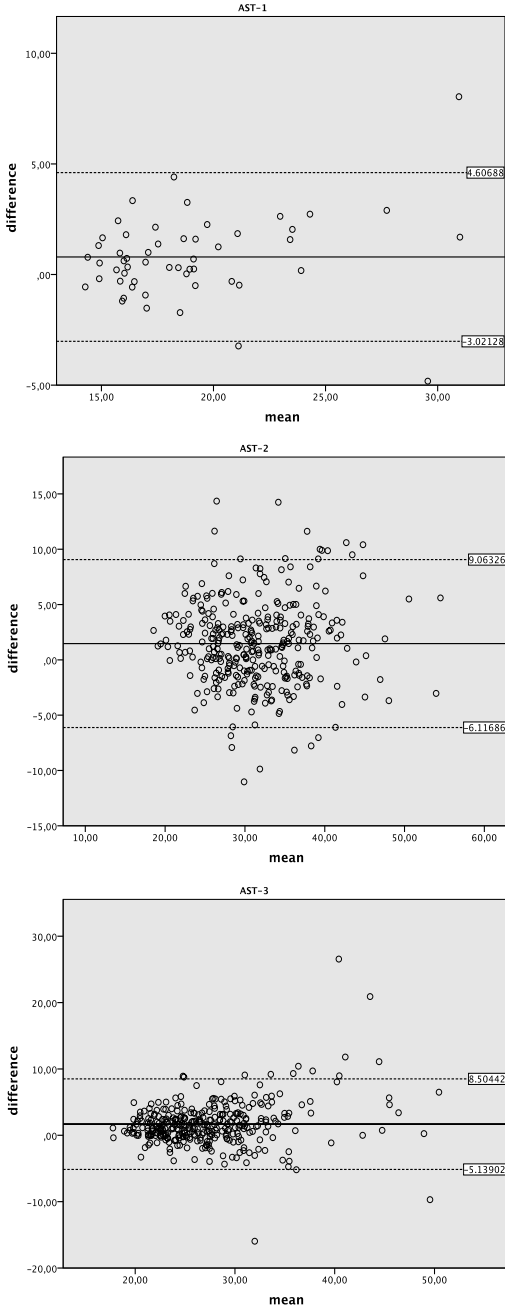
A high degree of test-retest reliability of the AST was found. The intraclass correlation coefficient between the first and second trial was 0.881 (95% CI: 0.780–0.934) for AST-1, 0.802 (95% CI: 0.717–0.858) for AST-2 and 0.800 (95% CI: 0.669–0.871) for AST-3.

**Table 3.2** Time to complete the Athletic Skills Tracks per track for each study.

		Study 1				Study 2	
		Trial 1 (baseline)		Trial 2 (2-3 weeks later)		Mean	(SD)
		Mean	(SD)	Mean	(SD)		
AST-1 (sec)	Boys	19.9	(4.9)	18.6	(3.9)	22.0	(4.6)
	Girls	18.7	(3.7)	18.4	(4.1)	21.9	(4.1)
	Total	19.3	(4.4)	18.5	(3.9)	22.0	(4.3)
AST-2 (sec)	Boys	31.2	(6.1)	29.6	(6.1)	30.3	(7.0)
	Girls	33.3	(6.5)	31.8	(6.4)	33.0	(6.8)
	Total	32.2	(6.4)	30.7	(6.3)	31.7	(7.0)
AST-3 (sec)	Boys	27.2	(6.1)	25.8	(5.6)	23.2	(4.0)
	Girls	29.4	(6.1)	27.4	(5.2)	27.8	(4.5)
	Total	28.3	(6.3)	26.6	(5.5)	25.4	(4.8)

AST-1= Athletic Skills Track 1; AST-2= Athletic Skills Track 2; AST-3= Athletic Skills Track 3.

The Bland and Altman plots for test-retest reliability of AST- 1 (mean = 0.79, [LoA] -3.02 and 4.60), AST-2 (mean = 1.47, [LoA] -6.12 and 9.06), and AST-3 (mean = 1.68, [LoA] -5.14 and 8.50) show that the differences between the test-retest do not change with increasing mean values. This indicates that there is no systematic bias (see Figure 3.4). In addition, the test-retest reliability difference (test 2 – test 1) and the mean of the trials [(trial 2 + trial 1)/2] shows no evidence of proportional bias between the two trials (AST-1: B = 0.108, P = 0.10; AST-2: B = 0.014, P = 0.69), except for AST-3 (B = 0.142, P = 0.00).



**Figure 3.4** Bland and Altman plots of AST-1, AST-2 and AST-3 (AST-1 = Athletic Skills Track 1; AST-2 = Athletic Skills Track 2; AST-3 = Athletic Skills Track 3).

The internal consistency of the AST, as represented by the Cronbach’s  $\alpha$ , was at or above the acceptable level of 0.70 (Cronbach, 1951) i.e., for AST-1:  $\alpha = 0.764$ , for AST-2:  $\alpha = 0.700$  and for AST-3:  $\alpha = 0.763$  (see also Table 3.3). Next, corrected item-total correlations were computed. All movement skill tasks in AST-1, AST-2 and AST-3 proved to be relevant for the reliability of the tracks. Table 3.3 shows that the Cronbach’s  $\alpha$  is lower if items are deleted except for the second split time (task “traveling jumps”) on AST-1. When examining the item redundancy of the AST in more detail, Table 3.3 shows a corrected item-total correlation of the split time of the movement skill tasks of AST-1 between 0.342 and 0.687, for AST-2 between 0.356 and 0.678, and for AST-3 between 0.407 and 0.627, indicating that the items measure the same construct (Everitt & Skrondal, 2010; Field, 2005). With respect to the concurrent validity of the AST, a moderate to high correlation was found between the time to complete the AST and the MQ KTK (i.e., AST-1:  $r = -0.747$ ,  $p = 0.01$ ; AST-2:  $r = -0.646$ ,  $p = 0.01$ ; and AST-3:  $r = -0.602$ ,  $p = 0.01$ ).



**Table 3.3** Internal consistency of the Athletic Skills Track.

	<b>Item</b>	<b>N</b>	<b>mean (sec)</b>	<b>(SD)</b>	<b>Corrected Item-total Correlation</b>	<b>Cronbach's Alpha If item deleted</b>
<b>AST-1</b>	<b>1</b>	39	6.6	(1.3)	.687	.663
	<b>2</b>	39	4.2	(1.2)	.342	.783
	<b>3</b>	39	3.4	(1.3)	.543	.719
	<b>4</b>	39	3.2	(0.7)	.675	.712
	<b>5</b>	39	4.4	(1.4)	.556	.718
	<b>Total</b>	39	22.0	(4.3)		.764
<b>AST-2</b>	<b>1</b>	72	6.9	(2.8)	.520	.651
	<b>2</b>	72	4.9	(1.3)	.678	.609
	<b>3</b>	72	4.7	(1.2)	.411	.674
	<b>4</b>	72	4.5	(1.7)	.390	.673
	<b>5</b>	72	3.5	(1.5)	.356	.683
	<b>6</b>	72	7.4	(2.0)	.415	.667
	<b>Total</b>	72	31.7	(7.0)		.700
<b>AST-3</b>	<b>1</b>	102	4.6	(1.5)	.627	.694
	<b>2</b>	102	3.8	(0.9)	.611	.711
	<b>3</b>	102	3.8	(0.7)	.446	.749
	<b>4</b>	102	3.5	(1.0)	.568	.714
	<b>5</b>	102	4.1	(1.1)	.407	.753
	<b>6</b>	102	5.6	(1.5)	.489	.743
	<b>Total</b>	102	25.4	(4.9)		.763

AST-1= Athletic Skills Track 1; AST-2= Athletic Skills Track 2; AST-3= Athletic Skills Track 3.

### 3.4 Discussion

This is the first study to examine the test-retest reliability, internal consistency, and concurrent validity of an age-related version of the AST in a PE setting.

The findings show that the Athletic Skills Track is a reliable and valid motor skill competence test that can be used to assess the FMS of 4- to 12-year old children in three age-bands in the PE setting. The results show a high level of test-retest reliability of the AST. Intraclass correlation coefficients of the three age-bands are all above 0.800. Bland and Altman plots show a high level of agreement between two trials of the AST completed within two weeks (Shrout & Fleiss, 1979) and the concurrent validity of the AST shows a moderate to high correlation with the KTK (Hinkle et al., 2003).

There is no sign of systematic or proportional bias for AST-1 and AST-2. Only AST-3 shows a proportional bias. When looking at the Bland and Altman plot of AST-3 the proportional bias might be caused by the outliers. The test – retest results are in line with Hoeboer et al. (2016) in which an ICC of 0.891 (95% CI: 0.870–0.908) was found between two trials that were completed within one day. The results are slightly lower than the test-retest reliability of the Canadian Agility and Movement Skill Assessment (CAMSA) (Lander, Morgan, Salmon, Logan, & Barnett, 2017; Longmuir et al., 2015). The test-retest reliability of the CAMSA was 0.91 in a sample of 34 adolescent girls (Lander et al., 2017).

The internal consistency of the AST is acceptable, with Cronbach's Alphas ranging from 0.700 up to 0.764 for the three age-bands (Cronbach, 1951). Except for the task "Traveling jumps" in AST-1, deleting tasks did not lead to higher Cronbach's Alphas.

The concurrent validity of the AST is decreasing with age (AST-1:  $r = -0.747$ ,  $p = 0.01$ ; AST-2:  $r = -0.646$ ,  $p = 0.01$ ; and AST-3:  $r = -0.602$ ,  $p = 0.01$ ), but the correlation coefficient is still above the threshold of a moderate to high correlation (Hinkle et al., 2003). The small decrease of the correlation coefficient could be explained by the fact that crucial development in the fundamental motor pattern period is related to further development of motor skill competence as shown in the mountain of motor development (Clark & Metcalfe, 2002; Seefeldt, 1980). Because the AST focusses on the fundamental motor pattern period AST-3 might not discriminate enough between children of different age groups from 9 till 12 years. Future research should focus on more profound research on different age groups of AST-3. Despite the decrease of the correlation, the correlation coefficients are higher than in the previous study on the AST in which the correlation between the time to complete a non-age-related version of the AST and the KTK ranged between  $r = -0.474$  and  $r = -0.502$  (Hoeboer et al., 2016). In

addition, the concurrent validity of the AST is within the range of the validity of other motor skill competence tests. For example, the correlation between the CAMSA and the Victorian FMS assessment in a sample of 34 adolescent girls was  $r = 0.680$  ( $P = 0.05$ ) (Lander et al., 2017).

In summary, the results show that the clinimetric quality of the age-related version of the AST was higher than the previous version of the AST (Hoeboer et al., 2016). In addition, the reliability and concurrent validity of the three age-bands of the AST are comparable with other, motor skill competence tests (Cools et al., 2009). However, most existing motor skill competence assessments are not feasible in the daily PE practice. In this study, it was investigated if it is possible to administer the AST during a regular PE lesson. A complete class of 24–30 children can be measured within one regular PE lesson of one hour. In addition, no special equipment is required for the AST. The only outcome measure is time which seems to be a reliable outcome. Another advantage of the AST is that the track can be used for children aged 4- to 12-years old, a critical phase in the development of children's motor competence (Clark, 2005).

Besides, rather than assessing isolated movements, the AST consists of a series of 5–7 concatenated FMS providing an assessment that more closely mimics how movement skills are performed in real life. Tidén et al. (2015) argued that one of the shortcomings of most motor skill competence test is the fact that they show low ecological validity. Since the skills in the AST are completed one after another and change according to the various constraints of the environment, the ecological validity of the AST might be better than that of other motor skill competence tests.

Despite the advantages of the AST it is important to keep the purpose of the assessment in mind and the reason for choosing one assessment over the other. Time, effort and level of experience required to execute a motor skill assessment contributes to the choice of the assessment (Logan et al., 2017). It seems that process (or a combination of process and product) orientated assessments are more time consuming than a product orientated assessment as the AST which only has time as outcome measurement (Cools et al., 2009). The AST gives an opportunity to screen complete school populations of children. To get a more detailed idea of the motor competence of a subgroup it might be comprehensive to add a follow-up process-orientated assessment. By including both types of assessments (product- and process orientated), a more detailed image of motor skill competence could yield (Logan et al., 2017).

This research has some limitations. The influence of children's BMI on the score on the AST needs to be studied, since previous research showed that childhood overweight results in poorer performances on a motor competence test (Vandorpe et al., 2011). In addition, sports participation was not controlled for. Although the track consisted of FMS rather than sport-specific tasks, participating in certain sports might have been advantageous. For example, a child participating in gymnastics might score better on the AST, because the skills in the track are closely related to certain gymnastics skills. Registration of sports participation and physical activity might provide insight in the ecological validity of the AST. Finally, a relatively small number of children completed AST-1 compared to AST-2 and AST-3. Therefore, the clinimetric results of this version of the test should be considered with some caution.

#### 3.4.1 Conclusions

In conclusion, the Athletic Skills Track seems to be a reliable and valid motor skill competence test that can be used to assess motor skill competence of 4- to 12-year children in a PE setting. The track measures motor skill competence of elementary school children in three age-bands. The AST provides a general indication of motor skill competence that can be compared to a reference norm.

By implementing the AST into the PE setting it becomes possible to compare motor skill competence of children, groups of children, schools and even regions in an objective way, just like periodic test for arithmetic, reading, grammar, etc. This could lead to more objectively insight of PE teachers in motor skill competence of children which could support PE teachers and other professionals to (re)define PE lessons and interventions with more focus on motor skill competence in the PE setting based on the test results of the AST.

### 3.5 References

- Altman, D. (1991). Inter-rater agreement. *Practical Statistics for Medical Research*, 5, 403–409.
- Barnett, L. M., Stodden, D., Cohen, K. E., Smith, J. J., Lubans, D. R., Lenoir, M., & Morgan, P. J. (2016). Fundamental movement skills: An important focus. *Journal of Teaching in Physical Education*, 35(3), 219–225. doi:10.1123/jtpe.2014-0209AQ5
- Barnett, L. M., Van Beurden, E., Morgan, P. J., Brooks, L. O., & Beard, J. R. (2009). Childhood motor skill proficiency as a predictor of adolescent physical activity. *Journal of Adolescent Health*, 44(3), 252–259. doi:10.1016/j.jadohealth.2008.07.004
- Bland, J. M., & Altman, D. (1986). Statistical methods for assessing agreement between two methods of clinical measurement. *The Lancet*, 327 (8476), 307–310. doi:10.1016/S0140-6736(86)90837-8
- Bland, J. M., & Altman, D. G. (1999). Measuring agreement in method comparison studies. *Statistical Methods in Medical Research*, 8(2), 135–160. doi:10.1177/096228029900800204
- Cale, L., Harris, J., & Chen, M. H. (2014). Monitoring health, activity and fitness in physical education: Its current and future state of health. *Sport, Education and Society*, 19(4), 376–397. doi:10.1080/13573322.2012.681298
- Clark, J. E. (2005). From the beginning: A developmental perspective on movement and mobility. *Quest*, 57(1), 37–45. doi:10.1080/00336297.2005.10491841
- Clark, J. E., & Metcalfe, J. S. (2002). The mountain of motor development: A metaphor. *Motor Development Research and Reviews*, 2, 163–190.
- Cools, W., De Martelaer, K., Samaey, C., & Andries, C. (2009). Movement skill assessment of typically developing preschool children: A review of seven movement skill assessment tools. *Journal of Sports Science & Medicine*, 8(2), 154–168.

Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*, 16(3), 297–334. doi:10.1007/BF02310555

D'Hondt, E., Deforche, B., Vaeyens, R., Vandorpe, B., Vandendriessche, J., Pion, J., Lenoir, M. (2011). Gross motor coordination in relation to weight status and age in 5-to 12-year-old boys and girls: A cross-sectional study. *International Journal of Pediatric Obesity*, 6(sup3), e556–e564. doi:10.3109/17477166.2010.500388

Everitt, B., & Skrondal, A. (Eds.). (2010). *The Cambridge dictionary of statistics* (4th ed.). New York, NY: Cambridge University Press.

Field, A. (2005). *Discovering statistics using SPSS*. London: SAGE Publications.

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

Hallal, P. C., Andersen, L. B., Bull, F. C., Guthold, R., Haskell, W., & Ekelund, U.; Lancet Physical Activity Series, Working Group. (2012). Global physical activity levels: Surveillance progress, pitfalls, and prospects. *The Lancet*, 380(9838), 247–257. doi:10.1016/S0140-6736(12)60646-1

Henderson, S. E., Sugden, D. A., & Barnett, A. L. (2007). *Movement assessment battery for children-2*. London: Harcourt Assessment.

Hinkle, D. E., Wiersma, W., & 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., & Savelsbergh, G. (2016). Validity of an athletic skills track among 6- to 12-year-old children. *Journal of Sports Sciences*, 34(21), 2095–2105.

Janssen, I., & LeBlanc, A. G. (2010). Systematic review of the health benefits of physical activity and fitness in school-aged children and youth. *The International Journal of Behavioral Nutrition and Physical Activity*, 7(1), 1–16. doi:10.1186/1479-5868-7-40

Kiphard, E. J., & Schilling, F. (2007). *Körperkoordinationstest für Kinder*. 2. Überarbeitete und ergänzte Auflage. Weinheim: Beltz Test GmbH.

Lander, N., Morgan, P. J., Salmon, J., & Barnett, L. M. (2016). Teachers' perceptions of a fundamental movement skill (FMS) assessment battery in a school setting. *Measurement in Physical Education and Exercise Science*, 20(1), 50–62. doi:10.1080/1091367X.2015.1095758

Lander, N., Morgan, P. J., Salmon, J., Logan, S. W., & Barnett, L. M. (2017). The reliability and validity of an authentic motor skill assessment tool for early adolescent girls in an Australian school setting. *Journal of Science and Medicine in Sport*, 20(6), 590–594. doi:10.1016/j.jsams.2016.11.007

Lenoir, M., Vandorpe, B., D'Hondt, E., Pion, J., Vandendriessche, J., Vaeyens, R., & Philippaerts, R. M. (2014). *KTK-NL Körperkoordinationstest für Kinder*. Herwerkte, gehernormeerde en vertaalde uitgave van de KTK voor het Nederlandstalig gebied (Vol. 1, pp. 80). Destelbergen: Universiteit Gent.

Lima, R. A., Pfeiffer, K., Larsen, L. R., Bugge, A., Moller, N. C., Anderson, L. B., & Stodden, D. F. (2017). Physical activity and motor competence present a positive reciprocal longitudinal relationship across childhood and early adolescence. *Journal of Physical Activity and Health*, 14(6), 1–21.

Lloyd, M., Saunders, T. J., Bremer, E., & Tremblay, M. S. (2014). Long-term importance of fundamental motor skills: A 20-year follow-up study. *Adapted Physical Activity Quarterly*, 31(1), 67–78. doi:10.1123/apaq.2013-0048

Logan, S. W., Barnett, L. M., Goodway, J. D., & Stodden, D. F. (2017). Comparison of performance on process- and product-oriented assessments of fundamental motor skills across childhood. *Journal of Sports Sciences*, 35(7), 634–641. doi:10.1080/02640414.2016.1183803

Longmuir, P. E., Boyer, C., Lloyd, M., Yang, Y., Boiarskaia, E., Zhu, W., & Tremblay, M. S. (2015). The canadian assessment of physical literacy: Methods for children in grades 4 to 6 (8 to 12 years). *BMC Public Health*, 15(1), 767. doi:10.1186/s12889-015-2106-6

Lubans, D. R., Morgan, P. J., Cliff, D. P., Barnett, L. M., & Okely, A. D. (2010). Fundamental movement skills in children and adolescents: Review of associated health benefits. *Sports Medicine*, 40(12), 1019–1035. doi:10.2165/11536850-000000000-00000

Miller, J., Vine, K., & Larkin, D. (2007). The relationship of process and product performance of the two-handed sidearm strike. *Physical Education & Sport Pedagogy*, 12(1), 61–76. doi:10.1080/17408980601060291

Okely, A. D., Booth, M. L., & Patterson, J. W. (2001). Relationship of cardiorespiratory endurance to fundamental movement skill proficiency among adolescents. *Pediatric Exercise Science*, 13(4), 380–391. doi:10.1123/pes.13.4.380

Pate, R. R., Pratt, M., Blair, S. N., Haskell, W. L., Macera, C. A., Bouchard, C., King, A. C. (1995). Physical activity and public health: A recommendation from the centers for disease control and prevention and the american college of sports medicine. *273*(5), 402–407. doi:10.1001/jama.1995.03520290054029

Robinson, L. E., Stodden, D. F., Barnett, L. M., Lopes, V. P., Logan, S. W., Rodrigues, L. P., & D'Hondt, E. (2015). Motor competence and its effect on positive developmental trajectories of health. *Sports Medicine*, 45(9), 1273–1284. doi:10.1007/s40279-015-0351-6

Runhaar, J., Collard, D. C., Singh, A. S., Kemper, H. C., Van Mechelen, W., & Chinapaw, M. (2010). Motor fitness in dutch youth: Differences over a 26- year period (1980–2006). *Journal of Science and Medicine in Sport/Sports Medicine Australia*, 13(3), 323–328. doi:10.1016/j.jsams.2009.04.006

Seefeldt, V. (1980). Developmental motor patterns: Implications for elementary school physical education. *Psychology of Motor Behavior and Sport*, 36(6), 314–323.



Shrout, P. E., & Fleiss, J. L. (1979). Intraclass correlations: Uses in assessing rater reliability. *Psychological Bulletin*, 86(2), 420–428. doi:10.1037/0033-2909.86.2.420

Stodden, D., Goodway, J., Langendorfer, S., Robertson, M., Rudisill, M., Garcia, C., & Garcia, L. (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

Tidén, A., Lundqvist, C., & Nyberg, M. (2015). Development and initial validation of the NyTid test: A movement assessment tool for compulsory school pupils. *Measurement in Physical Education and Exercise Science*, 19(1), 34–43. doi:10.1080/1091367X.2014.975228

Tremblay, M. S., Gray, C. E., Akinroye, K., Harrington, D. M., Katzmarzyk, P. T., Lambert, E. V., Tomkinson, G. (2014). Physical activity of children: A global matrix of grades comparing 15 countries. *Journal of Physical Activity and Health*, 11(s1), S113–S125. doi:10.1123/jpah.2014-0177

Turvey, M. T. (1990). Coordination. *American Psychologist*, 45(8), 938. doi:10.1037/0003-066X.45.8.938

Ulrich, D. (2000). TGMD 2—Test of gross motor development examiner's manual (Vol. 2). Austin, TX: PRO-ED.

Van Mechelen, W., Twisk, J. W., Post, G. B., Snel, J., & Kemper, H. C. (2000). Physical activity of young people: The Amsterdam longitudinal growth and health study. *Medicine & Science in Sports & Exercise*, 32(9), 1610–1616. doi:10.1097/00005768-200009000-00014

Vandorpe, B., Vandendriessche, J., Lefèvre, J., Pion, J., Vaeyens, R., Matthys, S. A., Lenoir, M. (2011). The KörperkoordinationsTest für kinder: Reference values and suitability for 6–12-year-old children in Flanders. *Scandinavian Journal of Medicine & Science in Sports*, 21(3), 378–388. doi:10.1111/sms.2011.21.issue-3

Wormhoudt, R., Teunissen, J. W., & Savelsbergh, G. J. (2012). Athletic skills model. Nieuwegein: Arko Sports Media.



# 4

## The Athletic Skills Track: Age- and gender-related normative values of a motor skills test for 4- to 12-year-old children

Hoeboer, J. J. A. A. M., Ongena, G., Krijger-Hombergen, M., Stolk, E., Savelsbergh, G. J. P., & de Vries, S. I. (2018).

*Journal of Science and Medicine in Sport, Volume 20, e76 - e77.*

**Objectives:** The aim of the presented study is to provide age- and gender-related normative values and MQ values for a motor skills test, the Athletic Skills Track, among 4- to 12-year-old children. **Methods:** In 2016, a total of 7977 Dutch children, 4036 boys (mean age 8.6 years, SD 2.1) and 3941 girls (mean age 8.6 years, SD 2.1), performed an age-related version of the Athletic Skills Track (AST). The AST is a track consisting of 5–7 fundamental movement skill tasks that should be completed as fast as possible. The children performed the test during a regular physical education (PE) lesson under the supervision of their own PE teacher. For each version of the AST (AST-1: n = 917; AST-2: n = 3947; AST-3: n = 3213) age- and gender-related reference centiles were derived from the gathered data using the Lambda, Mu, Sigma (LMS) method. **Results:** All children completed the AST within 60 s (mean 29.6 s, SD 7.7). An independent samples t- test showed that boys were significantly faster in completing the track than girls, except for the 4- year-old boys. Therefore, age- and gender-related reference centiles were derived. The reference curves demonstrate an almost linear decrease in time to complete AST-1 and AST-2 with increasing age. **Conclusions:** The present study provides age- and gender-related normative values and MQ values for the AST among 4- to 12-year-old Dutch children. With these normative values PE teachers can interpret children's performance on the AST.

## 4 The Athletic Skills Track: Age- and gender-related normative values of a motor skills test for 4- to 12-year-old children

### 4.1 Introduction

Physical literacy (PL) is gaining more and more attention in physical education (PE), physical activity (PA) and sports promotion worldwide. PL is the motivation, confidence, physical competence, knowledge and understanding to value and take responsibility for engagement in physical activities for life (Whitehead, 2013). PE is an important resource to develop PL during childhood (Cools, De Martelaer, Samaey, & Andries, 2009; Vandorpe et al., 2011). A high level of motor skill competence in youth is associated with several health-related parameters such as lower body mass index (D'Hondt et al., 2011; Okely, Booth, & Chey, 2004), better cardiorespiratory fitness (Okely, Booth, & Patterson, 2010), and higher PA levels in cross-sectional (Logan et al., 2014; Lubans, Morgan, Cliff, Barnett, & Okely, 2010; Wrotniak, Epstein, Dorn, Jones, & Kondilis, 2006) and longitudinal studies (Lloyd, Saunders, Bremer, & Tremblay, 2014). Motor skill competence is constrained by the level of fundamental movement skills (FMS) performance (Veldman, Jones, & Okely, 2016). FMS are associated with increased cognitive development (Best, 2010; Hillman, Erickson, & Hatfield, 2017), social development and language skills (Leonard & Hill, 2014). Furthermore, good FMS performance is positively associated with self-esteem (Piek, Baynam, & Barrett, 2006) and reduced levels of anxiety (Skinner & Piek, 2001).

The crucial development of FMS is imbedded in the development of motor skill competence as shown in the mountain of motor development (Clark & Metcalfe, 2002). Although motor skill development is a nonlinear, self-organizing process that is driven by task, environment and organism (Newell, 1986), this metaphor of motor development shows six periods of development that characterizes most typically developing individuals (Clark & Metcalfe, 2002).

Since children's FMS and PA levels seem to have decreased in the last decades (Hallal et al., 2012), there is an urgency to increase our understanding of the development of FMS and to develop effective strategies to support children in obtaining an optimal level of PL. Therefore, we need valid motor skill competence tests. The urgency to increase our understanding of motor skill development is also recognised by PE teachers who are willing to monitor motor skill competence of children more objectively (Lander, Morgan, Salmon, & Barnett, 2016).

In 2016, researchers (Hoeboer et al., 2016) have developed a robust and feasible motor skill competence test, the Athletic Skills Track (AST) that PE teachers can use to assess children's FMS in an objective way. The AST is fundamentally different than conventional motor competence tests. The AST is a track that consists of a series of 5 to 7 detached, rather than isolated, FMS based on coordinative abilities (Wormhoudt, Teunissen, & Savelsbergh, 2012) (e.g. coupling, spatial orientation, and balance ability) (Rudd et al., 2015; Tidén, Lundqvist, & Nyberg, 2015). Previous studies have shown that the AST is a reliable, valid and feasible assessment tool to assess FMS among children from 4 to 12 years old in the PE setting (Hoeboer, Krijger-Hombergen, Savelsbergh, & De Vries, 2017; Hoeboer et al., 2016). In the first study, it was shown that it was possible to measure one class of approximately 24 children in a one hour lesson. All children completed the track within 60 seconds. The concurrent validity of the AST was moderate to high; i.e., the correlation coefficient between the time to complete the AST and age- and gender-specific motor quotients of the Körperkoordination-Test für Kinder (KTK) ranged between  $r = -0.469$  and  $r = -0.767$  ( $p < 0.05$ ) (Hoeboer et al., 2016). In the second study three age-related versions of AST have been developed. AST-1 has been developed for 4- to 6-year-old children, AST-2 for 6- to 9-year-old children and AST-3 for 9- to 12-year-old children. The test-retest reliability of the AST proved to be high (AST-1: ICC = 0.881, AST-2: ICC = 0.802, AST-3: ICC = 0.800). The internal consistency was above the acceptable level of Cronbach's  $\alpha > 0.70$  (AST-1 = 0.764, AST-2 = 0.700, AST-3 = 0.763). The concurrent validity between the AST and the KTK were even higher than in the first study (AST-1:  $r = -0.747$ ,  $p = 0.01$ ; AST-2:  $r = -0.646$ ,  $p = 0.01$ ; and AST-3:  $r = -0.602$ ,  $p = 0.01$ ) (Hoeboer et al., 2017).

Although the AST appears to be a promising objective tool to assess children's motor skill competence in the PE setting, the lack of normative values limits the ability to interpret children's performance on the AST. Producing reference norms and Motor Quotient (MQ) outcomes generates a scoring system that describes individuals global motor competence in meaningful categories, and allows comparison with results from other schools, region, countries and studies (Toftegaard-Stoeckel, Groenfeldt, & Andersen, 2010). Therefore, the objective of the presented cross-sectional study was to provide age- and gender-related normative values and MQ values for the AST for Dutch children from 4 to 12 years old.

## 4.2 Method

The research design was set up to reach a large group of children. Children aged 4- to 12-years were recruited from 86 primary schools in The Hague region, the Netherlands. Informed consent was obtained from the parents or guardians of the children after they were given written information about the purpose and nature of the study. After the informed consent was obtained, 8,458 children completed an age-specific version of the Athletic Skills Track (AST-1, AST-2 or AST-3) under the supervision of their own PE teacher.

The age-specific versions of the AST consist of a series of fundamental movement skill tasks (AST-1:  $n = 5$ , AST-2:  $n = 7$ , and AST-3:  $n = 7$ ) to be completed as fast as possible (see appendix 2 - Maps of AST-1, AST-2 and AST-3) (Hoeboer et al., 2017). The only outcome measurement is time to complete the track. In all three tracks the same FMS are tested, with the difficulty of the tasks ascending from AST-1 over AST-2 to AST-3. The AST was completed during a regular PE lesson between March and April 2016. In this lesson the measurement was conducted in a gymnasium that was separated into three sections (see appendix 2 - map of the gymnasium into three sections). In two sections the children received an autonomous assignment from their PE teacher. In the third part the AST was conducted. The PE teacher performed the track once after which the children performed three try-out trials. After having received feedback from the PE teacher during the three try-out trials the children performed one measurement trial independently. The PE teacher measured the time to complete the track using a stopwatch and registered it in Excel.

All 86 PE teachers who participated in this study were trained according to the AST protocol by the principal investigator of this study. The protocol provides guidelines about the AST to be used per age group. It also provides guidelines about how to deal with adverse events, such as falling or cheating. The study protocol was approved by the Ethical Committee of the Faculty of Human Movement Sciences, VU University Amsterdam, The Netherlands (ECB 2015–31). Data analysis was performed with the Statistical Package for the Social Sciences (SPSS version 24.0, 64-bits edition, SPSS Inc, Chicago, Illinois). Of the 8,458 children who were allowed by their parents or guardians to participate in the study, 7,977 children met the inclusion criteria for data analysis (age: between 4 and 12 years; all data complete: Age, Gender, AST-track number and AST-time).

All data were expressed as Mean, Standard Deviation (SD), and range. The normality of the data was investigated by analysing the normal distribution in relation to the medians. In addition, histograms were plotted. Differences between boys and girls were examined per track

with independent samples t-tests and with a one-way ANOVA it was investigated if the tracks were able to distinguish between age groups within the track. Age- and gender-related reference centiles were derived per track using the Lambda, Mu, Sigma (LMS) method as introduced by Cole (1992) using R GAMLSS packages.

Based on the reference centiles of the AST, MQ values were calculated following the example of the Körperkoordination-Test für Kinder (KTK) (Kiphard & Schilling, 2007). The MQ (mean = 100; SD = ± 15) gives an indication of children's level of motor giftedness, ranging from "gifted children" to "children with motor dysfunctions" (Kiphard & Schilling, 2007). The norms for the MQ of the KTK are based on the performance of 1228 normally developing German children (1974). The MQ score is standardized by age and gender. As stated by Kiphard and Schilling (2007) in a normal population, a MQ score below 85 represents a motor performance level that is considered as problematic. The normalised value was scaled with 100 as the average value based on the 50th percentile. The other cut-off points are based on the 10th, 25th, 75th, 90th percentiles using the following formula:  $MQ = (50th\ percentile\ AST-x / time\ AST-x) * 100$ .

#### 4.3 Results

The final sample consisted of 7,977 4- to 12-year old children (4036 boys and 3941 girls). The mean age of the boys was 8.6 years (SD 2.1) and the mean age of the girls was 8.6 years (SD 2.1). Appendix 3 - Descriptive statistics of respondents' characteristics, provides descriptive statistics of the final sample as well as the Dutch population (Central Dutch Foundation for Statistics [CBS]) (CBS Statline, 2017). In our sample, the ratio of boys to girls was approximately equal, which does not significantly deviate from the Dutch youth population. Considering age, the sample does not fully reflect the Dutch population in the lower age group. This is because not all primary schools in The Netherlands provide PE lessons for children at the ages of 4 and 5 years. Hence, we consider the sample to be representative for Dutch children attaining PE lessons in terms of gender and age.

As shown in Table 4.1 all children completed the AST within 60 seconds. On average, boys completed AST-1 in 25.3 seconds ± 7.1, AST-2 in 30.6 seconds ± 7.3, and AST-3 in 27.0 seconds ± 6.9. Girls completed AST-1 in 27.4 ± 7.9 seconds, AST-2 in 33.0 seconds ± 7.9, and AST-3 in 29.1 seconds ± 6.8.

**Table 4.1** Descriptive statistics for each group of participants.

Track	N	Mean	(± SD)	95% CI	
				Lower	Upper
AST-1, Boys, age 4	78	32.3	(7.6)	30.6	34.0
AST-1, Boys, age 5	200	25.3	(7.0)	24.4	26.3
AST-1, Boys, age 6	199	22.6	(5.0)	21.9	23.3
<b>AST-1, Boys, total</b>	<b>477</b>	<b>25.3</b>	<b>(7.1)</b>	<b>24.7</b>	<b>26.0</b>
AST-1, Girls age 4	93	34.2	(8.8)	32.4	35.9
AST-1, Girls age 5	171	27.1	(7.1)	26.1	28.2
AST-1, Girls, age 6	176	24.0	(5.4)	23.2	24.8
<b>AST-1, Girls, total</b>	<b>440</b>	<b>27.4</b>	<b>(7.9)</b>	<b>26.6</b>	<b>28.1</b>
AST-2, Boys, age 6	222	34.6	(7.3)	33.6	35.6
<b>AST-2, Boys, age 7</b>	<b>630</b>	<b>32.5</b>	<b>(7.3)</b>	<b>31.9</b>	<b>33.1</b>
AST-2, Boys, age 8	633	29.6	(7.0)	29.1	30.1
AST-2, Boys, age 9	485	27.4	(5.8)	26.9	27.9
<b>AST-2, Boys, total</b>	<b>1970</b>	<b>30.6</b>	<b>(7.3)</b>	<b>30.2</b>	<b>30.9</b>
AST-2, Girls, age 6	235	38.4	(8.0)	37.4	39.4
AST-2, Girls, age 7	636	35.1	(7.8)	34.5	35.8
AST-2, Girls, age 8	622	30.9	(6.7)	30.4	31.4
AST-2, Girls, age 9	384	29.7	(6.7)	29.0	30.3
<b>AST-2, Girls, total</b>	<b>1877</b>	<b>33.0</b>	<b>(7.9)</b>	<b>32.7</b>	<b>33.4</b>
AST-3, Boys, age 9	140	27.1	(5.9)	26.1	28.1
AST-3, Boys, age 10	506	27.5	(6.7)	27.0	28.1
AST-3, Boys, age 11	510	27.3	(7.4)	26.7	28.0
AST-3, Boys, age 12	433	25.9	(6.5)	25.2	26.5
<b>AST-3, Boys, total</b>	<b>1589</b>	<b>27.0</b>	<b>(6.9)</b>	<b>26.6</b>	<b>27.3</b>
AST-3, Girls, age 9	163	29.4	(6.2)	28.5	30.4
AST-3, Girls, age 10	500	30.4	(7.0)	29.8	31.0
AST-3, Girls, age 11	535	28.7	(6.9)	28.2	29.3
AST-3, Girls, age 12	426	28.0	(6.6)	27.3	28.6
<b>AST-3, Girls, total</b>	<b>1624</b>	<b>29.1</b>	<b>(6.8)</b>	<b>28.8</b>	<b>29.5</b>

AST-1 = Athletic Skills Track-1, AST-2 = Athletic Skills Track-2 and AST-3 = Athletic Skills Track-3.



An independent samples t-test showed a significant difference between boys and girls in time to complete the track (see Table 4.2). Boys were significantly faster than girls (on average between 1.3 up to 3.8 seconds), except for the 4-year-olds on AST-1.

**Table 4.2** Results of an independent-samples t-test between boys and girls.

Track	Boys	Girls	T	P
AST-1, age 4	32.3 (± 7.6) (N=78)	34.2 (± 8.8) (N=93)	-1.457	0.147
AST-1, age 5	25.3 (± 6.9) (N=200)	27.1 (± 7.1) (N=171)	-2.457	0.014*
AST-1, age 6	22.6 (± 5.0) (N=199)	24.0 (± 5.4) (N=176)	-2.642	0.009*
AST-2, age 6	34.6 (± 7.3) (N=222)	38.4 (± 8.0) (N=235)	-5.261	0.000*
AST-2, age 7	32.5 (± 7.3) (N=630)	35.1 (± 7.8) (N=636)	-6.236	0.000*
<b>AST-2, age 8</b>	<b>29.6 (± 7.0)</b> <b>(N=633)</b>	<b>30.9 (± 6.7)</b> <b>(N=622)</b>	<b>-3.388</b>	<b>0.001*</b>
AST-2, age 9	27.4 (± 5.8) (N=485)	29.7 (± 6.7) (N=384)	-5.147	0.000*
AST-3, age 9	27.1 (± 5.9) (N=140)	29.4 (± 6.1) (N=163)	-3.260	0.001*
AST-3, age 10	27.6 (± 6.7) (N=506)	30.7 (± 7.0) (N=500)	-6.532	0.000*
AST-3, age 11	27.3 (± 7.4) (N=510)	28.7 (± 6.7) (N=535)	-3.278	0.001*
AST-3, age 12	25.9 (± 6.6) (N=433)	28.0 (± 6.6) (N=426)	-4.791	0.000*

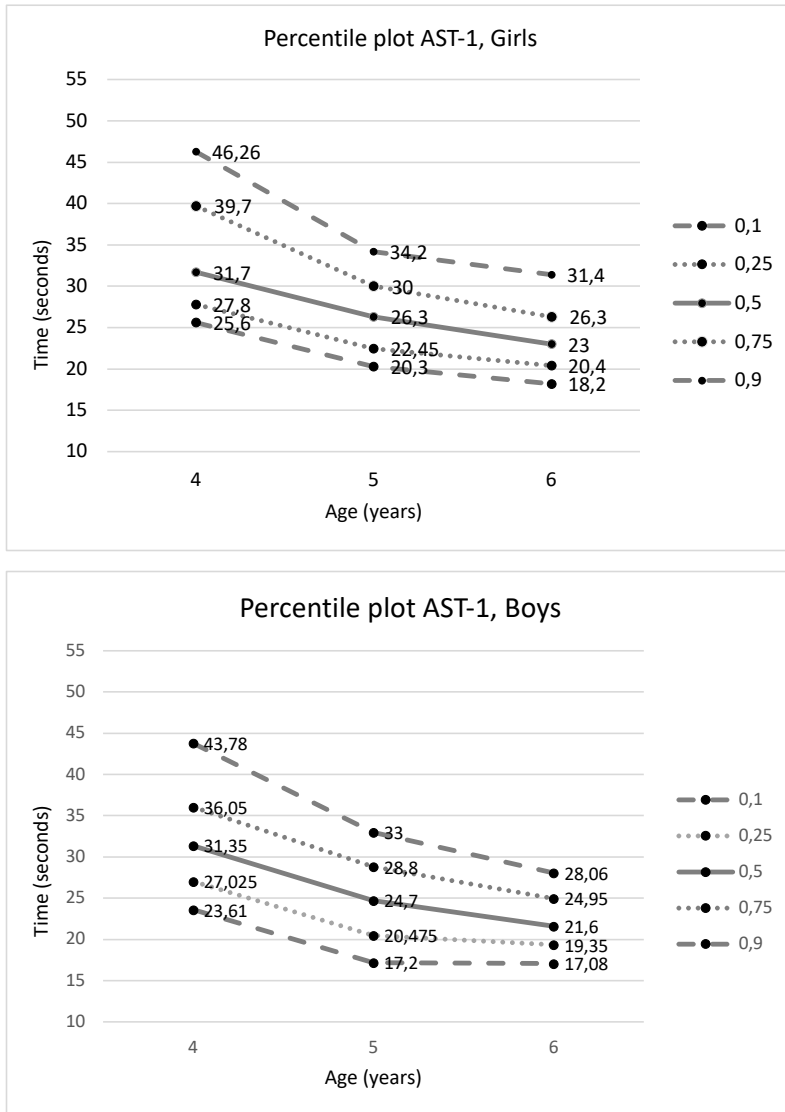
AST-1 = Athletic Skills Track-1, AST-2 = Athletic Skills Track-2 and AST-3 = Athletic Skills Track-3.

A one-way ANOVA revealed that, both for boys and for girls, the tracks were able to distinguish between age groups within the track. The time scores on AST-1, AST-2 and AST-3 were significantly different between the age groups (AST- 1 girls: Welch's F (65.412) = 3138.491, P < 0.01; AST-2 girls: Welch's F (105.520) = 5596.357, P < 0.01; AST-3 girls: Welch's F (10.702) = 485.279, P < 0.01; AST- 1 boys: Welch's F (65.685) = 2639.745, P < 0.01; AST-2 boys: Welch's F (80.036) = 3782.352, P < 0.01; AST-3 boys: Welch's F

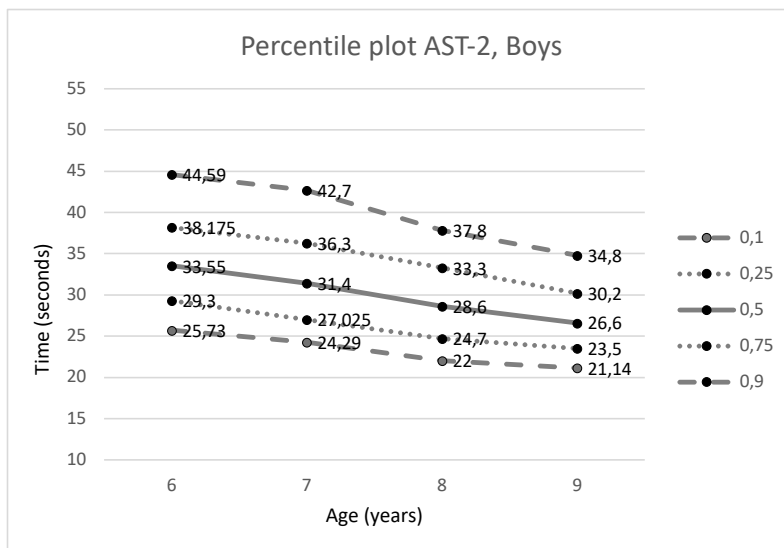
(5.634) = 262.541,  $P < 0.01$ ). Bonferroni post hoc analysis revealed that the differences between the age groups within a track were all significant ( $P < 0.01$ ) except for the difference between some age groups for AST-3 (AST-3 girls: mean difference 9-10 year olds: -1.002 seconds, 95% CI: -2.606 - 0.603; mean difference 9-11 year olds: 0.673 seconds, 95% CI: -0.919-2.264; mean difference 9-12 year olds: 1.409 seconds, 95% CI: -0.233-3.043; mean difference 11-12 year olds: 0.732 seconds, 95% CI: -0.423-1.887; and AST-3 boys: mean difference 9-10 year olds: -0.448 seconds, 95% CI: -2.170 - 1.274; mean difference 9-11 year olds: -0.182 seconds, 95% CI: -1.902-1.539; mean difference 9-12 year olds: 1.274 seconds, 95% CI: -0.479-3.028; mean difference 10-11 year olds: 0.266 seconds, 95% CI: -0.865-1.398).

As can be seen in Table 4.2, the results of 6-year olds on AST-1 cannot be compared with the time of (other) 6-year old children to complete AST-2. It took 6-year-old boys  $34.6 \pm 7.3$  seconds to complete AST-2 compared to  $22.6 \pm 5.0$  seconds for AST-1. Similar findings are shown for girls. In the transition period between AST-2 and AST-3 smaller differences are found. The 9-year-old boys and girls needed more or less the same time to complete AST-2 compared to AST-3 (boys:  $27.4 \pm 5.8$  versus  $27.1 \pm 5.9$  seconds for AST-2 and AST-3 respectively; girls:  $29.7 \pm 6.7$  versus  $29.4 \pm 6.2$  seconds for AST-2 and AST-3 respectively). Based on the significant differences between boys and girls, reference centiles were calculated per gender and track using the Lambda, Mu, Sigma (LMS) method 27 based on the 10th, 25th, 50th, 75th and 90th percentile. As shown in Figure 4.1 and 4.2 the reference curves demonstrate an almost linear decrease in time to complete AST-1 and AST-2 with increasing age for boys and girls. The time to complete AST-3 remains relatively stable among 9- to 12-year-old boys and girls (see Figure 4.3).

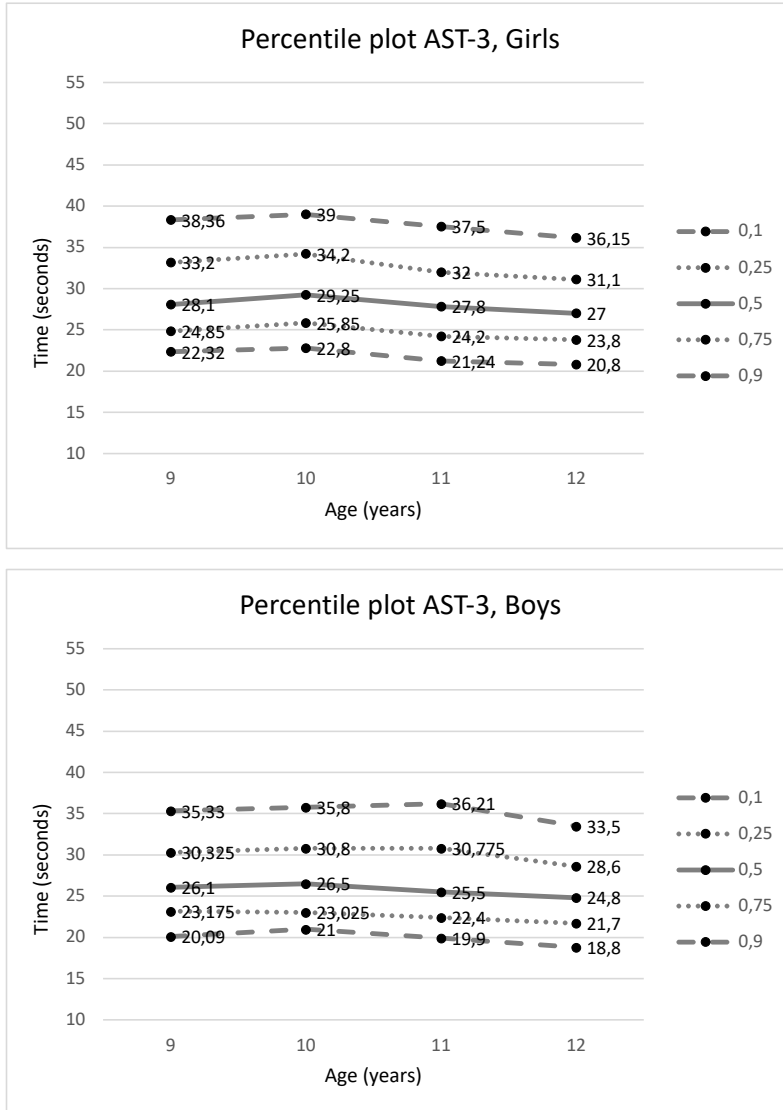
To ease the interpretation of the outcomes of the AST the reference curves have been converted into MQ values (see Table 4.3 – MQ values for five categories of motor giftedness). With the formula  $MQ = (50\text{th percentile AST-}x / \text{time AST-}x) * 100$  and the table PE teachers can calculate and interpret each child's MQ value based on the time to complete the track, the 50th percentile of the track, the age and the gender of the child. For example, a 7-year-old boy finishes AST-2 in 29.5 seconds. The MQ of this boy is:  $MQ = 31.4$  (50th percentile of AST-2) / 29.5 seconds (time of the boy to complete AST-2) \* 100 =  $MQ = 31.4 / 29.5 * 100 = 106$ . This child is normally motor gifted since his MQ value lies between 87 and 116 (see appendix 3 – MQ values for five categories of motor giftedness).



**Figure 4.1** Gender specific percentile plots for AST-1. The percentile plots are based on the 10<sup>th</sup>-25<sup>th</sup>-50<sup>th</sup>-75<sup>th</sup>-90<sup>th</sup> percentile. AST-1 = Athletic Skills Track-1.



**Figure 4.2** Gender specific percentile plots for AST-2. The percentile plots are based on the 10th-25th-50th-75th-90th percentile. AST-2 = Athletic Skills Track-2.



**Figure 4.3** Gender specific percentile plots for AST-3. The percentile plots are based on the 10<sup>th</sup>-25<sup>th</sup>-50<sup>th</sup>-75<sup>th</sup>-90<sup>th</sup> percentile. AST-3 = Athletic Skills Track-3.

**Table 4.3** Distribution of Motor Quotients (MQ) of the total samples based on formula:
$$\text{MQ person } x = (0,5^{\text{th}} \text{ percentile AST-}x / \text{time AST-}x) * 100.$$

Track	Severe motor disorder	Moderate motor disorder	Normal motor gifted	Good motor gifted	High motor gifted
AST-1, Boys, age 4	72	87	100	116	133
AST-1, Boys, age 5	75	86	100	121	144
AST-1, Boys, age 6	77	87	100	112	126
AST-1, Girls age 4	69	80	100	114	124
AST-1, Girls age 5	77	88	100	117	130
AST-1, Girls, age 6	73	87	100	113	126
AST-2, Boys, age 6	75	88	100	115	130
AST-2, Boys, age 7	74	87	100	116	129
AST-2, Boys, age 8	76	86	100	116	130
AST-2, Boys, age 9	76	88	100	113	126
AST-2, Girls, age 6	77	86	100	118	133
AST-2, Girls, age 7	74	85	100	114	130
AST-2, Girls, age 8	75	88	100	116	131
AST-2, Girls, age 9	76	88	100	116	134
AST-3, Boys, age 9	74	86	100	113	130
AST-3, Boys, age 10	74	86	100	115	126
AST-3, Boys, age 11	70	83	100	114	128
AST-3, Boys, age 12	74	87	100	114	132
AST-3, Girls, age 9	73	85	100	113	126
AST-3, Girls, age 10	75	86	100	113	128
AST-3, Girls, age 11	74	87	100	115	131
AST-3, Girls, age 12	75	87	100	113	130
<b>Average</b>	<b>74</b>	<b>86</b>	<b>100</b>	<b>115</b>	<b>130</b>
<b>KTK cut of points</b>	<b>70</b>	<b>85</b>	<b>100</b>	<b>115</b>	<b>130</b>

AST-1 = Athletic Skills Track-1, AST-2 = Athletic Skills Track-2 and AST-3 = Athletic Skills Track-3.

#### 4.4 Discussion

In this study age- and gender-related normative values have been developed based on the time to complete AST-1, AST-2 and AST-3 for children in the age of 4- to 12-years old.

The results show that boys are significantly faster on all the tracks except for the 4-year-old boys (boys:  $32.3 \pm 7.6$  seconds versus girls:  $34.2 \pm 8.8$  seconds;  $t = -1.457$ ,  $p = 0.147$ ). The gender difference that we have found is in line with other motor competence tests such as the KTK (Kiphard & Schilling, 2007).

Besides gender differences, age differences have been found as well. As shown in the reference curves an almost linear decrease in time to complete AST-1 and AST-2 can be seen with increasing age. The time to complete AST-3 remains relatively stable among 9- to 12-year-old boys and girls. This phenomenon is in line with the theoretical metaphor: “the mountain of motor development” developed by Clark and Metcalfe (2002). This metaphor of motor development characterizes most typically developing individuals in six periods (Clark & Metcalfe, 2002). In the fundamental patterns period, fundamental motor skills are further elaborated into the building blocks of later context specific motor skills. This fundamental patterns period lasts for most children until the age of 7 to 8 years old (Clark & Metcalfe, 2002). The development of the fundamental movement skills in this period can be seen in the reference centiles of AST-1 and AST-2 as these children are in the age of developing those fundamental building blocks. The relatively stable line of the reference centiles of AST-3 can be explained by the fact that the children are at the end of the fundamental pattern period and at the beginning of the following, context-specific, period. This can also explain the lack of difference in the time to complete AST-2 and AST-3 for 9-year old children. Nine-year-old boys and girls needed approximately the same time to complete AST-2 as AST-3 (AST-2: boys:  $27.4 \pm 5.8$  seconds, girls:  $29.7 \pm 6.7$  seconds versus AST-3: boys:  $27.1 \pm 5.9$  seconds, girls:  $29.4 \pm 6.2$  seconds). In this context-specific period children have established their basic motor repertoire and they begin to apply the fundamental patterns towards a variety of tasks and environmental contexts (Clark & Metcalfe, 2002). The transitions between AST-1, AST-2 and AST-3 shows that the 6-year-old boys ( $22.6 \pm 5.0$ ) and girls ( $24.0 \pm 5.4$ ) on AST-1 are faster than the 6-year-old boys ( $34.6 \pm 7.3$ ) and girls ( $38.4 \pm 8.0$ ) on AST-2. This is in line with the fact that AST-1 consists of 5 skills and AST-2 of 7 skills.

Besides reference curves based on time to complete the tracks, MQ values have been calculated for the AST. This eases the interpretation of AST outcomes in daily PE practice. The development of MQ values can be seen as a scoring system that classifies individual scores

into meaningful global categories and allows comparison with results from other studies (Toftegaard-Stoekel et al., 2010).

According to the classification of Kiphard and Schilling (2007), children with an MQ value between 86 and 115 are considered as having normal gross motor coordination, children scoring between 71 and 85 as having a moderate gross motor coordination disorder and children scoring 70 or less as having a severe gross motor coordination disorder. Children scoring between 116 and 130 are considered as having good motor coordination and children scoring above 130 as having a high motor coordination. Based on the reference centiles in this study approximately the same cut-off points are found for AST-1, AST-2 and AST-3. These findings are in line with the study on the reference values of the KTK (Vandorpe et al., 2011). This study has some limitations. First, the participants of this study are all children in the The Hague region. This might have influenced the normative values presented because the data was collected within the urbanity of The Hague. A sub-urban population might show other results. Secondly, the data was collected among a sample of children who all received 2 PE lessons per week by qualified PE teachers. In other parts of the Netherlands this is not the case. On average, 54% of the primary school in the Netherlands have a qualified PE teacher to provide the PE lessons (Reijgersberg & Lucassen, 2013). This might have influenced the results because children who receive PE lessons from a qualified PE teacher might have better motor skills. In future studies it could be interesting to examine the reference centiles in other regions of the Netherlands or other countries.

Thirdly, the reference curves presented in this study are based on the age in years. This means that all children with the same birth year are presented in the reference curves as the same age. It would be more precise if the reference curves are based on the year and month of birth.

#### 4.4.1 Conclusion

In conclusion, the AST is a reliable, valid and feasible assessment tool to assess FMS among children from 4- to 12 in the PE setting (Hoeboer et al., 2017; Hoeboer et al., 2016). With the presented age- and gender-related normative values and MQ values for the AST PE teachers can compare the motor skill competence of individual children with the motor skill competence of their peers or with their previous scores. In addition, the motor skill competence of groups of children, schools, neighbourhoods, regions, and even countries can be compared to the presented age- and gender-related normative values.



#### 4.4.2 Practical implications

- PE teachers can use the AST to assess FMS of 4-to 12-year old children in the PE setting;
- PE teachers can use the AST normative values to interpret the level of motor competence of individual children and groups of children;
- The MQ values eases the interpretation of AST outcomes in daily PE practice.

## 4.5 References

Best, J. R. (2010). Effects of physical activity on children's executive function: Contributions of experimental research on aerobic exercise. *Developmental Review*, 30(4), 331-351.

CBS Statline. (2017). Population dynamics: Birth death and migration per region. From <http://statline.cbs.nl/StatWeb/publication/?VW=T&DM=SLEN&PA=37259eng&LA=EN>

Clark, J. E., & Metcalfe, J. S. (2002). The mountain of motor development: A metaphor. *Motor Development Research and Reviews*, 2, 163-190.

Cole, T. J., & Green, P. J. (1992). Smoothing reference centile curves: The LMS method and penalized likelihood. *Statistics in Medicine*, 11(10), 1305-1319.

Cools, W., De Martelaer, K., Samaey, C., & Andries, C. (2009). Movement skill assessment of typically developing preschool children: A review of seven movement skill assessment tools. *Journal of Sports Science and Medicine*, 8(2), 154-168.

D'Hondt, E., Deforche, B., Vaeyens, R., Vandorpe, B., Vandendriessche, J., Pion, J., Lenoir, M. (2011). Gross motor coordination in relation to weight status and age in 5-to 12-year-old boys and girls: A cross-sectional study. *International Journal of Pediatric Obesity*, 6(sup3), e556-564.

Hallal, P. C., Andersen, L. B., Bull, F. C., Guthold, R., Haskell, W., Ekelund, U., & Lancet Physical Activity Series, Working Group. (2012). Global physical activity levels: Surveillance progress, pitfalls, and prospects. *Lancet* (London, England), 380(9838), 247-257.

Hillman, C. H., Erickson, K. I., & Hatfield, B. D. (2017). Run for your life! childhood physical activity effects on brain and cognition. *Kinesiology Review*, 6(1), 12-21.

Hoeboer, J., Krijger-Hombergen, M., Savelsbergh, G., & De Vries, S. (2017). Reliability and concurrent validity of a motor skill competence test among 4- to 12-year old children. *Journal of Sports Sciences*, 1-7.

Hoehoer, J., De Vries, S., Krijger-Hombergen, M., Wormhoudt, R., Drent, A., Krabben, K., & Savelsbergh, G. (2016). Validity of an athletic skills track among 6-to 12-year-old children. *Journal of Sports Sciences*, 1-11.

Kiphard, E. J., & Schilling, F. (2007). *Körperkoordinationstest für kinder: KTK Beltz-Test*.

Lander, N., Morgan, P. J., Salmon, J., & Barnett, L. M. (2016). Teachers' perceptions of a fundamental movement skill (FMS) assessment battery in a school setting. *Measurement in Physical Education and Exercise Science*, 20(1), 50-62.

Leonard, H. C., & Hill, E. L. (2014). Review: The impact of motor development on typical and atypical social cognition and language: A systematic review. *Child and Adolescent Mental Health*, 19(3), 163-170.

Lloyd, M., Saunders, T. J., Bremer, E., & Tremblay, M. S. (2014). Long-term importance of fundamental motor skills: A 20-year follow-up study. *Adapted Physical Activity Quarterly*, 31(1), 67-78.

Logan, S. W., Robinson, L. E., Getchell, N., Webster, E. K., Liang, L., & Golden, D. (2014). Relationship between motor competence and physical activity: A systematic review. *Research Quarterly for Exercise and Sport*, 85(S1), A14.

Lubans, D. R., Morgan, P. J., Cliff, D. P., Barnett, L. M., & Okely, A. D. (2010). Fundamental movement skills in children and adolescents. *Sports Medicine*, 40(12), 1019-1035.

Newell, K. M. (1986). Constraints on the development of coordination. *Motor Development in Children: Aspects of Coordination and Control*, 34, 341-360.

Okely, A. D., Booth, M. L., & Chey, T. (2004). Relationships between body composition and fundamental movement skills among children and adolescents. *Research Quarterly for Exercise and Sport*, 75(3), 238-247.

Okely, A. D., Booth, M. L., & Patterson, J. W. (2001). Relationship of cardiorespiratory endurance to fundamental movement skill proficiency among adolescents. *Pediatric Exercise Science*, 13(4), 380-391.

Piek, J. P., Baynam, G. B., & Barrett, N. C. (2006). The relationship between fine and gross motor ability, self-perceptions and self-worth in children and adolescents. *Human Movement Science*, 25(1), 65-75.

Reijgersberg, N., & Lucassen, J. (2013). *Nulmeting bewegingsonderwijs: Onderzoek naar de organisatie van het bewegingsonderwijs in het primair onderwijs*. Utrecht: Mulder Instituut.

Rudd, J. R., Barnett, L. M., Butson, M. L., Farrow, D., Berry, J., & Polman, R. C. (2015). Fundamental movement skills are more than run, throw and catch: The role of stability skills. *PloS One*, 10(10), e0140224.

Skinner, R. A., & Piek, J. P. (2001). Psychosocial implications of poor motor coordination in children and adolescents. *Human Movement Science*, 20(1), 73-94.

Tidén, A., Lundqvist, C., & Nyberg, M. (2015). Development and initial validation of the NyTid test: A movement assessment tool for compulsory school pupils. *Measurement in Physical Education and Exercise Science*, 19(1), 34-43.

Toftgaard-Stoekel, J., Groenfeldt, V., & Andersen, L. B. (2010). Children's self-perceived bodily competencies and associations with motor skills, body mass index, teachers' evaluations, and parents' concerns. *Journal of Sports Sciences*, 28(12), 1369-1375.

Vandorpe, B., Vandendriessche, J., Lefevre, J., Pion, J., Vaeyens, R., Matthys, S., Lenoir, M. (2011). The KörperkoordinationsTest für kinder: Reference values and suitability for 6-12-year-old children in flanders. *Scandinavian Journal of Medicine & Science in Sports*, 21(3), 378-88.

Veldman, S. L., Jones, R. A., & Okely, A. D. (2016). Efficacy of gross motor skill interventions in young children: An updated systematic review. *BMJ Open Sport & Exercise Medicine*, 2(1), e000067.

Whitehead, M. (2013). Definition of physical literacy and clarification of related issues. *ICSSPE Bulletin*, 65(1.2).

Wormhoudt, R., Teunissen, J. W., & Savelsbergh, G. J. (2012). *Athletic skills model*. Nieuwegein: Arko Sports Media.

Wrotniak, B. H., Epstein, L. H., Dorn, J. M., Jones, K. E., & Kondilis, V. A. (2006). The relationship between motor proficiency and physical activity in children. *Pediatrics*, 118(6), e1758-65. doi:118/6/e1758 [pii]



# Children's enjoyment of a motor skill test in Physical Education

Hoeboer, J., De Vries, S., Mast, D., & Savelsbergh, G. (2017).

*Journal of Physical Education Research, Volume 4, Issue IV, 01-16.*

A novel motor skill test, the Athletic Skills Track (AST) was developed in 2016, to assess motor skill competence of children in the PE setting. The main purpose of this study was to examine children's enjoyment when being tested with the AST in a regular PE lesson. The study was conducted among 239 children, aged 4-to-12-years-old. All children completed the AST, where after they rated their enjoyment using a Smileyometer. One week later, 131 children completed another motor skill test, the Körperkoordinationstest für Kinder (KTK), where after they were asked to rank their enjoyment with the AST, the KTK, measurements of body height and weight, and a periodical cognitive test, using a Funsorter. The majority (98%) of the children rated their enjoyment of the AST as good to brilliant. 76% of the children ranked the AST as the most enjoyable in relation to the other three tests. No significant differences were found in enjoyment of the AST between boys and girls, nor between children with different motor competence levels. Most the children enjoyed performing the AST in a regular PE lesson. Future studies should examine the effect of feedback and repeated testing on children's enjoyment.

## 5 Children's enjoyment of a motor skill test in Physical Education

### 5.1 Introduction

The decrease in health-related parameters among children, such as their motor skill competence, physical activity (PA) level, and sedentary behaviour, has become a worldwide concern in the last decades (Barnett, Van Beurden, Morgan, Brooks, & Beard, 2009; Biddle & Asare, 2011; Chen, Mason, Hypnar, & Bennett, 2016; D'Hondt et al., 2014; Dollman, Norton, & Norton, 2005). Various government policies have foregrounded schools to be instrumental in addressing these issues (Cale & Harris, 2013). This has changed the role of schools and physical education (PE) over the last decade (Cale & Harris, 2006; Shephard & Trudeau, 2010; Webb, Quennerstedt, & Öhman, 2008). Since then, various initiatives have been integrated into the school and PE curricula to promote public health (Green, 2014). Recently, a shift in focus can be observed in PE from health- and fitness-related goals to aspects of physical literacy. Physical literacy is "the motivation, confidence, physical competence, knowledge and understanding to value and take responsibility for engagement in physical activities for life" (Whitehead, 2013). The concerns on children's health and the embracement of physical literacy in PE has led to more health-related testing in the PE setting (Sum et al., 2016). The focus on physical literacy has also renewed the attention for children's motor skill competence, a basic element of physical literacy (Whitehead, 2013). According to Wrotniak, Epstein, Dorn, Jones and Kondilis (2006) motor skills competence is a crucial component for children's engagement in physical activity.

Children's motor skill competence can be measured with several assessment tools. Cools et al. (2009) reviewed such tools and stated that, in general, the tests are not very feasible in a PE setting. It takes at least 20 minutes to measure one individual child. Furthermore, special test materials and extensive knowledge of the test protocols are required to conduct the tests. Besides the burden for PE teachers, some researchers also warn that monitoring all kinds of health-related parameters within PE might be counterproductive to the promotion of PA, especially for the children most in need of encouragement (Cale & Harris, 2009; Naughton, Carlson, & Greene, 2006). The measurement can be unpleasant or even embarrassing for children because it measures the children who are at risk in a confronting manner (Cale & Harris, 2009; Stanec, 2009). However, testing in PE can be valuable. For example, it can contribute in supporting a physically active and healthy lifestyle and in enhancing the PE curriculum (Lloyd, Colley, & Tremblay, 2010; Rowland, 2007; Silverman, Keating, & Phillips, 2008). Cale et al. (2014) have described a list of recommendations for PE teachers how to



approach testing in PE to make it enjoyable for children. The most practical recommendations are: 1) make sure that the testing takes place in a regular context (PE lesson). 2) make the testing enjoyable for all children. 3) do not focus on the outcome of the measurements, and 4) carefully select an assessment tool that takes account of all children, also the children who need feedback on motor skill competence, PA and healthy lifestyle the most (Cale et al., 2014). These recommendations are in line with the positive pedagogical approaches to testing in PE of Silverman et al. (2008). They claim that testing should not be used in an isolated manner without an educational purpose.

In 2016, a novel motor skill test has been developed for 4- to 12-year old children that was especially designed to fit into the context of PE (Hoeboer et al., 2016). This test is called the Athletic Skills Track (AST). The main purpose of this study was to examine children's enjoyment when being tested with the AST in a regular PE lesson. We also examined whether boys have another level of enjoyment than girls when performing the AST and whether children with poorer motor skills have another enjoyment level of the AST than children with better motor skills. A secondary purpose of this study was to examine how children enjoy the AST compared to other tests that children are exposed to during their primary school period.

## 5.2 Methods

### 5.2.1 Setting and subjects

This study was conducted at a random selection of two internship schools from the The Hague University of Applied Sciences in The Netherlands. Both primary schools were located in the The Hague city area and were willing to participate in the study. All parents of the children from grade 1 to 8 (4-12 years old) received an information letter with details about the purpose and nature of the study, where after informed consent was obtained from the parents or guardians of all children. Ninety-three percent of the children ( $n = 239$ ; 136 boys and 103 girls) were allowed by their parents or guardians to participate in the study. The study protocol was approved by the Ethical Committee of the Faculty of Human Movement Sciences, VU University Amsterdam, The Netherlands (ECB 2015-31).

### 5.2.2 Measurements

Testing was spread out over a two-week time-period in February 2016. All children conducted the AST in the first measurement week, where after their enjoyment of the test was measured using a Smileyometer (Read, 2008). The AST was conducted during a regular PE lesson in a

separate section of the gym by two research assistants (fourth year PE students of the The Hague University of Applied Sciences) who had been trained in conducting the tests according to the protocol (Hoeboer et al., 2016) in two meetings.

To investigate the second research question, one week later 131 randomly selected children completed a more clinical motor skill test, the Körperkoordinationstest für Kinder (KTK) during a regular PE lesson. The KTK was conducted by the same two research assistants who were also trained in conducting the KTK according to the protocol (Lenoir et al., 2014; Vandorpe et al., 2011). When the children finished the KTK they were asked to rank their enjoyment of the AST, the KTK, measurements of body height and weight, and a periodical cognitive test called CITO, using a Funsorter (Risden, Hanna, & Kanerva, 1997).

Both the anthropometric measurements and the CITO test are executed once every year at the participating schools. The anthropometric measurements were conducted between January 2016 and February 2016 during a regular PE lesson by children's PE teacher in a separated room near the gym. The CITO test was executed in January 2016, under the supervision of children's group teacher in the classroom. In the following paragraphs each test will be described in more detail.

#### *5.2.2.1 Athletic Skills Track*

All children completed the Athletic Skills Track (AST). The AST is a motor skill test that has been developed to assess fundamental movement skills (FMS) among 4- to 12-year-old children in a PE setting (Hoeboer et al., 2016). The test was executed according to the test manual (see appendix 2). The AST contains three age-related tracks consisting of a series of 5 to 7 concatenated FMS (locomotive and stability skills) (AST-1:  $n = 5$ ; AST-2:  $n = 7$ ; and AST-3:  $n = 7$ ) to be completed as fast as possible. In all three tracks the same FMS are tested, but the difficulty of the tasks is ascending from AST-1 to AST-3. AST-1 has been designed for the youngest children in the age of 4 to 6 years and consists of the following skills: 1) Walking, 2) Traveling jumps, 3) Alligator crawl, 4) Slaloming, and 5) Clambering. Children between 6 and 9 years performed AST-2. In this track the following skills had to be performed: 1) Walking, 2) Traveling jumps, 3) Hopscotch, 4) Alligator crawl (backwards), 5) Running (backwards), 6) Pencil roll, and 7) Clambering. Children in the age of 9 to 12 years completed AST-3 consisting of: 1) Walking (backwards), 2) Traveling jumps, 3) Bunny hopping, 4) Alligator crawl (backwards), 5) Slaloming (backwards), 6) Forward role, and 7) Clambering (see supplementary material: AST schematically displayed). Previous studies have shown that the AST is a valid and feasible motor skill test to assess children's FMS in a PE setting

(Hoeboer et al., 2016). The AST was performed during a regular PE lesson in 1/3rd of the gymnasium. In the other 2/3rd of the gymnasium the PE lesson was conducted as usual under the supervision of the PE teacher. The children practiced the track three times after an example given by the research assistant and then had an individual test attempt where time to complete the track was measured by a research assistant with a stopwatch. For all children, it was their first experience with the test. It took less than 1 minute per child to complete the AST.

#### 5.2.2.2 *Körperkoordinationstest für Kinder*

Children's enjoyment of the AST was compared to their enjoyment of a more clinical motor skill test, the KTK (Kiphard & Schilling, 2007). The children performed this test also for the first time. This test is not especially designed for the PE setting. The KTK consists of four subtests: (1) Walking backwards three times along each of three balance beams (3 m length; 6, 4.5, and 3 cm width, respectively; 5 cm height); (2) Moving across the floor in 20 seconds by stepping from one plate (25 cm × 25 cm × 5.7 cm) to the next by transferring the first plate, stepping on it, etc.; (3) Jumping from one leg over an increasing pile of pillows (60 cm × 20 cm × 5 cm each) after a short run-up; and (4) Jumping laterally as many times as possible over a wooden slat (60 cm × 4 cm × 2 cm) in 15 seconds. The test protocol for the Dutch language area was followed (Lenoir et al., 2014; Vandorpe et al., 2011). It took ± 25 minutes per child to complete the KTK.

#### 5.2.2.3 *Anthropometric measurements*

Children's enjoyment of the AST was also compared to their enjoyment of the measurement of their body height and weight. Children's body height and weight were measured by their PE teacher using a standard procedure of the municipality health services. Body weight was measured with a standard digital scale (Seca 877). Body height was measured using a standard measurement scale in millimeters. The children were individually measured in a separate space during a PE lesson. They were measured on bare feet in their sportswear (shorts and t-shirt). The results of the anthropometric measurements are not shown because they were not used for the purpose for this study.

#### 5.2.2.4 *CITO test*

To compare children's enjoyment of the AST to their enjoyment of an a-physical test they are familiar with in the school setting, the CITO test was included in the research design. The CITO

is a periodical cognitive test developed by the National Institute for Educational Measurement (CITO). The test gives an indication of the general proficiency level in cognitive development consisting of four subjects, i.e., Language, Math, Word orientation and Study skills. Children perform this test twice a year in the primary school period in a regular classroom setting under the supervision of their classroom teacher. All children of a class execute the test in the same timeframe. For the purpose of this study children were asked to remember their last experience with the CITO test in January 2016. The outcomes of the CITO are not shown because these results were not used for the purpose for this study.

#### 5.2.2.5 Smileyometer

In order to measure children's enjoyment of the AST a Smileyometer was used. The Smileyometer has been developed to measure children's enjoyment (Read, 2008). It is based on the Funometer, a 'thermometer' with a vertical bar representing the amount of fun (Read, MacFarlane, & Casey, 2002). The Smileyometer uses a pictorial representation of enjoyment ranging from 'awful' (1) to 'brilliant' (5) on a Likert scale as shown in Figure 5.1.

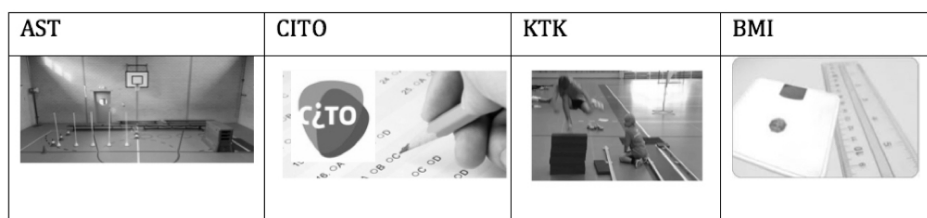
After the children finished the AST children were asked to, individually, select one smiley that fitted their enjoyment of the AST best, in the dressing room. Similar scales have been used in previous studies, for example in a study on the management of postoperative pain. Children have been presented with pain faces before and after surgery (Bosenberg, Thomas, Lopez, Kokinsky, & Larsson, 2003).



Figure 5.1 Smileyometer used to record children's enjoyment with the AST.

### 5.2.2.6 Funsorter

After the children finished the KTK ( $n = 131$ ) they were asked to rank their enjoyment of the AST, the KTK, measurements of body height and weight, and the CITO, using a Funsorter (Risdén, Hanna, & Kanerva, 1997). The purpose of the Funsorter is to rank a series of connected activities in order to establish which is most enjoyable (Sim, MacFarlane, & Read, 2006). After a short explanation, the children were asked to rank the four tests using four pictograms with an image of the tests. The children were asked to put the four loose pictograms in order from least enjoyable (1) to most enjoyable (4) to execute. The funsorter was conducted individually in the dressing room (see Figure 5.2: Funsorter).



**Figure 5.2** Funsorter (AST = Athletic Skills Track, CITO = a cognitive test, KTK = Körperkoordinationstest für Kinder, BMI = Anthropometric measurements).

### 5.2.3 Data analyses

Of the 239 children who were allowed by their parents or guardians to participate in the study all children met the inclusion criteria (age: between 4 and 12 years old). All the children completed the AST and the Smileyometer. A subgroup of 131 children also completed the KTK and the Funsorter. The outcome measurement of the AST is time to complete the track. The faster the child finishes the track the higher the motor skill competence (Hoeboer et al., 2016). The raw test scores on the KTK were converted into age- and gender-specific motor quotients (KTK MQ) using the test protocol for the Dutch language area (Lenoir et al., 2014). Children were then classified into five categories of motor giftedness based on the classification of Kiphard and Schilling (2007). Children with a MQ value between 86 and 115 are considered as having a normal gross motor coordination (NMG), children with a MQ value between 71 and 85 as having a moderate gross motor coordination disorder (MMD) and children scoring 70 or less as having a severe gross motor coordination disorder (SMD). Children scoring between 116 and 130 are considered as having a good motor coordination (GMG) and children scoring 131 and higher as high motor gifted (HMG).

Descriptive statistics were performed to characterize the sample and to report children's outcomes on the AST, the Smileyometer and the Funsorter. To analyze a normal distribution of the outcome parameters, histograms were plotted and the kurtosis and skewness values for each of the outcome parameters were assessed. Since the outcome measurement of the Smileyometer and the Funsorter were not normally distributed non-parametric tests were used. First, differences in enjoyment of the children performing AST-1, AST-2 or AST-3 were tested using a Somers d test. Secondly, differences between boys and girls in their enjoyment of the AST were tested non-parametrically using an Independent Samples Mann-Witney U test. The influence of motor giftedness on children's enjoyment was tested with an Independent Samples Kruskal-Wallis Test. Next, the Funsorter results were tested non-parametrically with the Related-Samples Friedman's Two-Way Analysis of variance by Ranks to examine if the enjoyment of the four tests differed significantly from each other. All statistical analyses were performed using IBM SPSS 23.0 64-bit edition. Values were considered statistically significant at  $P < 0.05$ .

### 5.3 Results

In total, 239 4- to 12-year old children (136 boys and 103 girls) completed the AST. Of this total sample, 131 children also completed the KTK (80 boys and 103 girls). The children's age and score on the AST and the KTK are shown in Table 5.1. On average, it took about  $22.5 \pm 4.0$  seconds to complete AST-1,  $26.1 \pm 4.6$  seconds to complete AST-2, and  $24.1 \pm 4.5$  seconds for AST-3. The average KTK Motor Quotient was  $90 \pm 11$ .

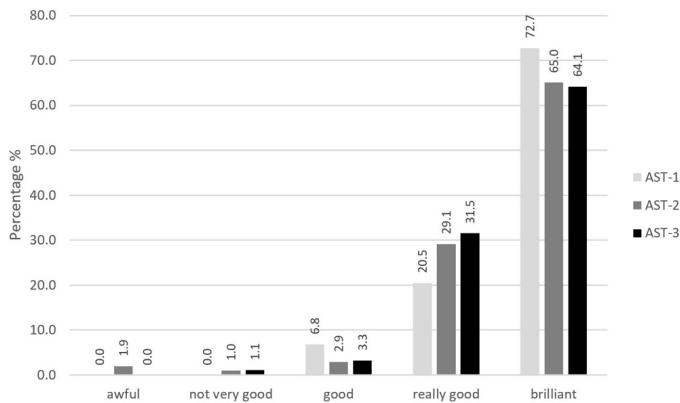
#### 5.3.1 Children's enjoyment of the AST

All children reported their enjoyment of the AST using the Smileyometer. The majority of the children (98%) indicated that the AST was enjoyable. They enjoyed the AST really good to brilliant (Figure 5.3). Of the 239 children two children reported that the AST was awful. There was no significant difference in enjoyment between children that completed AST 1, 2 and 3 (Somers d = 0.036,  $p = 0.489$ ).

**Table 5.1** Descriptive statistics of children's age and score on AST and KTK.

		N	Mean	( $\pm$ SD)	95% CI	
					Lower	Upper
Age (years)	Boys	136	8	2	7.92	8.64
	Girls	103	9	2	8.09	8.92
	Total	239	8	2	7.41	8.20
AST-1 (sec)	Boys	29	22.2	4.3	20.57	23.86
	Girls	15	23.1	3.3	21.27	24.89
	Total	44	22.5	4.0	21.29	23.71
AST-2 (sec)	Boys	60	25.7	4.5	24.55	26.88
	Girls	43	26.6	4.7	25.10	28.02
	Total	103	26.1	4.6	25.10	26.97
AST-3 (sec)	Boys	47	24.6	3.9	23.41	25.71
	Girls	45	23.8	5.1	22.27	25.33
	Total	92	24.1	4.5	23.25	25.13
KTK (MQ)	Boys	80	89	11	86.88	91.89
	Girls	51	90	11	86.67	92.79
	Total	131	90	11	87.61	91.43

AST-1 = Athletic Skills Track-1, AST-2 = Athletic Skills Track-2, AST-3 = Athletic Skills Track-3 and KTK = Körperkoordinationstest für Kinder.



**Figure 5.3** Level of reported enjoyment per Athletic Skills Track according to the Smileyometer .AST-1 N=44, AST-2 N=103, AST-3 N=92.

An Independent Samples Mann-Whitney U test was conducted to determine if there were differences in Smileyometer scores for boys and girls. Distributions of the Smileyometer scores were similar, as assessed by visual inspection. Median Smileyometer scores were not significantly different between boys and girls ( $U = 6.67$ ,  $z = -0.753$ ,  $p = 0.452$ ).

With an Independent Samples Kruskal-Wallis H Test the differences in enjoyment measured with the Smileyometer was investigated between children with different motor skill levels (MQ categories). Distributions of enjoyment measured with the Smileyometer were not similar for all groups (SMD, MMD, NMG, GMG, HMG) as assessed by visual inspection of a boxplot. However, the distribution of the Smileyometer scores is not significantly different across all categories of the MQ classification ( $X^2(3) = 1.623, p=0.654$ ) (see Figure 5.4. frequencies of reported enjoyment).

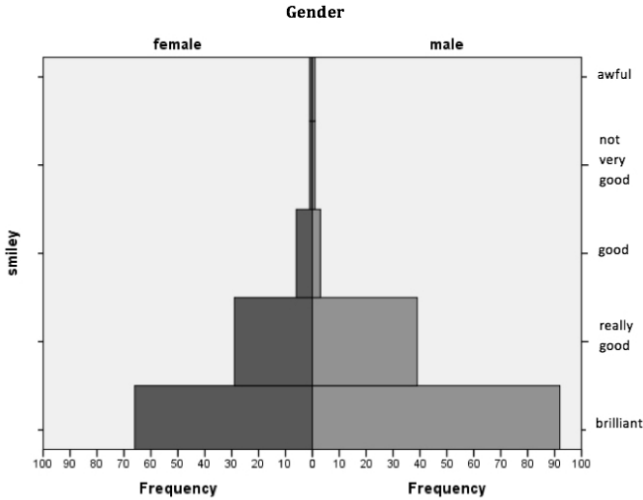


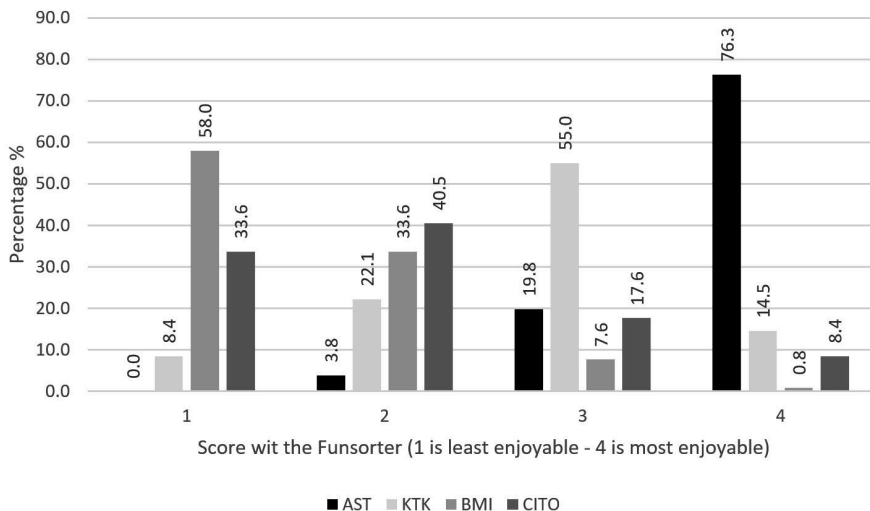
Figure 5.4 Frequencies of reported enjoyment

5.3.2 Children’s enjoyment of the AST related to other tests

In Figure 5.5 the results of the Funsorter are systematically displayed. It shows that the AST was never ranked as the least fun and 76% as most fun.

According to the Related-Samples Friedman’s Two-Way Analysis of variance by Ranks test the ‘fun’ scores were not evenly distributed across the four tests ( $X^2(3) = 220.409, p = 0.05$ ). Post hoc pairwise analysis revealed significant differences between AST (Mdn = 4.0), anthropometric measurements (Mdn = 1.0), CITO (Mdn = 2.0), and KTK (Mdn = 3.0) ( $p = 0.05$ ).





**Figure 5.5** Level of enjoyment during the AST, KTK, anthropometric measurements and the CITO-test as reported with the Funsorter (N=131)

#### 5.4 Discussion

As the AST was developed to assess children's FMS in a regular PE context taking account of all children, the main purpose of this study was to examine children's enjoyment of the AST. Other purposes were to examine if boys have another opinion than girls about the AST and to investigate if differences in motor giftedness resulted in another rating of the AST. A secondary purpose of this study was to examine how children enjoy the AST compared to other tests that children are exposed to during their primary school period.

Most children (98%) were positive about the AST. In addition, 76% of the children ranked the AST as the most fun to perform in relation to three other tests. No significant differences in enjoyment were found between age groups (AST-1, AST-2 and AST-3), nor between gender or children with different levels of motor giftedness. These findings are in line with the experience of PE teachers with the AST. In 2017, 86 primary schools in the The Hague region have used the AST to assess the FMS of 4- to 12-year old children (Hoeboer et al., 2017). Although data about PE teachers' experience was not gathered systematically, they were positive about the AST and did not report negative enjoyment of children with the AST. The results are also in line with our expectations because we took into account the recommendations of Cale et al. (2014). To our knowledge, this is the first study looking into enjoyment in relation to motor competence testing. Therefore, it is difficult to compare our findings with previous studies on motor competence assessment liking.

Although the Smileyometer has proved to be an easy and quick to use tool in this and previous studies (Tomlinson, von Baeyer, Stinson, & Sung, 2010) and it has shown to be a reliable and valid self-report scale for children in other studies, for example to indicate pain (Wong, 2001), the Smileyometer has some limitations that should be taken into account when interpreting the results of our study. First, social desirability cannot be ruled out when using a Smileyometer (King & Bruner, 2000). Perhaps that children have the tendency to respond in a manner that will be viewed favorably by others. This tendency poses a social desirability bias that should be considered, especially with the 'like'-culture nowadays. In addition, the Smileyometer was only used for the AST. Children might be as positive about the KTK, or the other two tests when using this 5-item Likert scale.

Secondly, there is inconclusive evidence about the age children are able to use the Smileyometer. A previous research shows that the Smileyometer is a reliable measurement tool for children older than 4 years old (Zaman, Abeele, & De Grooff, 2013). Other researchers on the other hand, claim that the Smileyometer is not very useful for children younger than 10 years old because the variability of the responses is very low. Children younger than 10 years old tend to choose the highest (most positive) score (Read, 2008).

Thirdly, one can question the sensitivity of 5-item Likert scale to assess differences between boys and girls or children with different levels of motor giftedness in children's enjoyment. However, in a review that compared different faces scales to indicate pain differences between boys and girls have been found in responsiveness, reliability and validity (Tomlinson et al., 2010).

Another limitation of this study that should be mentioned is the use of pictograms to rank the enjoyment of the four tests. The influence of the visual representation of the four tests depicted on the Funsorter cards cannot be ruled out. Perhaps photographs, or pictures with a child on it leads to a more positive test rating compared to more neutral or abstract pictures of the test. In future studies, the visual representation of the tests depicted on the cards should be at a similar level of iconicity, to avoid different interpretation of identical pictures.

The different time intervals between performing the four tests and rating them might have had an influence on the test outcomes, due to differences in retrieval of experiences after different time intervals. When children have a bad experience, the memories of that event tend to be stronger than with a positive experience (Baumeister, Bratslavsky, Finkenauer, & Vohs, 2001). In future research, similar time intervals should be used between the tests and ranking of the tests.

In the future, it might also be interesting to look at children's perceived motor competence in relation to their enjoyment of a motor competence test. According to Barnett et al. (2009) the perceived motor skill competence might have an important role in measuring assessment liking. Although children with different levels of - objectively measured - motor giftedness did not rate their enjoyment of the AST significantly different, children with a big difference between their own perceived level of motor skill competence and their objective level of motor skill competence might enjoy the AST differently.

Further research could also investigate children's enjoyment of the AST on the long term. The novelty of the AST could be causing a distortion in measuring its user experience. Novelty effect occurs when individuals who participate in research respond to a novel situation differently in the context of the study than they would in the real world (Gravetter & Forzano, 2003). When the AST loses its novelty, it might also lose its appeal. Furthermore, the effect of providing feedback and the way this is done to the children about their performance on the AST on assessment liking should be studied.

#### 5.4.1 Conclusions

In conclusion, the results of this study show that the majority of the 4- to 12-year old children enjoyed conducting the AST in a regular PE lesson. Their enjoyment is not dependent of age, gender or objective measured level of motor giftedness.

These findings, together with the good internal consistency, high test-retest reliability (range: 0.800 (95% CI: 0.669-0.871) - 0.881 (95% CI: 0.780-0.934)) and moderate to good concurrent validity when correlated to the Motor Quotient (MQ) of the KTK (range:  $r = -0.502$  ( $P < 0.01$ ) -  $r = -0.767$  ( $P < 0.01$ )) as shown in previous studies, strengthen the use of the AST in the PE setting ((Hoeboer et al., 2016).

## 5.5 References

- Barnett, L.M., van Beurden, E., Morgan, P.J., Brooks, L.O., & Beard, J.R. (2009). Childhood motor skill proficiency as a predictor of adolescent physical activity. *Journal of Adolescent Health, 44*(3), 252-259.
- Baumeister, R.F., Bratslavsky, E., Finkenauer, C., & Vohs, K.D. (2001). Bad is stronger than good. *Review of General Psychology, 5*(4), 323.
- Biddle, S.J., & Asare, M. (2011). Physical activity and mental health in children and adolescents: A review of reviews. *British Journal of Sports Medicine, 45*(11), 886-895.
- Bosenberg, A., Thomas, J., Lopez, T., Kokinsky, E., & Larsson, L. (2003). Validation of a six-graded faces scale for evaluation of postoperative pain in children. *Pediatric Anesthesia, 13*(8), 708-713.
- Cale, L. & Harris, J. (2006). School-based physical activity interventions: Effectiveness, trends, issues, implications and recommendations for practice. *Sport, Education and Society, 11*(4), 401-420.
- Cale, L. & Harris, J. (2009). Fitness testing in physical education-a misdirected effort in promoting healthy lifestyles and physical activity? *Physical Education and Sport Pedagogy, 14*(1), 89-108.
- Cale, L. & Harris, J. (2013). 'Every child (of every size) matters' in physical education! physical education's role in childhood obesity. *Sport, Education and Society, 18*(4), 433-452.
- Cale, L., Harris, J., & Chen, M. H. (2014). Monitoring health, activity and fitness in physical education: Its current and future state of health. *Sport, Education and Society, 19*(4), 376-397.
- Chen, W., Mason, S., Hypnar, A., & Bennett, A. (2016). Assessing motor skill competency in elementary school students: A three-year study. *Journal of Sports Science & Medicine, 15*(1), 102.
- Cools, W., De Martelaer, K., Samaey, C., & Andries, C. (2009). Movement skill assessment of typically developing preschool children: A review of seven movement skill assessment tools. *Journal of Sports Science & Medicine, 8*(2), 154-168.

D'Hondt, E., Deforche, B., Gentier, I., Verstuyf, J., Vaeyens, R., De Bourdeaudhuij, I., Lenoir, M. (2014). A longitudinal study of gross motor coordination and weight status in children. *Obesity* (Silver Spring, Md.), 22(6), 1505-1511.

Dollman, J., Norton, K., & Norton, L. (2005). Evidence for secular trends in children's physical activity behaviour. *British Journal of Sports Medicine*, 39(12), 892-897.

Gravetter, F.J. & Forzano, L.B. (2003). *Research methods for the behavioral sciences* Thomson. Cengage Learning, Boston.

Green, K. (2014). Mission impossible? reflecting upon the relationship between physical education, youth sport and lifelong participation. *Sport, Education and Society*, 19(4), 357-375.

Hoeboer, J., De Vries, S., Krijger-Hombergen, M., Wormhoudt, R., Drent, A., Krabben, K., & Savelsbergh, G. (2016). Validity of an athletic skills track among 6-to 12-year-old children. *Journal of Sports Sciences*, 1-11.

Hoeboer, J., Krijger-Hombergen, M., Savelsbergh, G., & De Vries, S. (2017). Reliability and concurrent validity of a motor skill competence test among 4-to 12-year old children. *Journal of Sports Sciences*, 1-7.

Hoeboer, J. J. A. A. M., Ongena, G., Krijger-Hombergen, M., Stolk, E., Savelsbergh, G. J. P., & de Vries, S. I. (2018). The Athletic Skills Track: Age-and gender-related normative values of a motor skills test for 4-to 12-year-old children. *Journal of Science and Medicine in Sport*.

King, M.F., & Bruner, G.C. (2000). Social desirability bias: A neglected aspect of validity testing. *Psychology and Marketing*, 17(2), 79-103.

Kiphard, E. J., & Schilling, F. (2007). *Körperkoordinationstest für kinder: KTK Beltz-Test*.

Lenoir, M., Vandorpe, B., Hondt, E., Pion, J., Vandendriessche, J., Vaeyens, R., & Philippaerts, R. (2014). D' & KTK-NL körperkoordinationstest für kinder. *Herwerkte Gehernormeerde En Vertaalde Uitgave Van De KTK Voor Het Nederlandstalig Gebied Destelbergen Universiteit Gent*, (volume 1, pp. 80).

Lloyd, M., Colley, R.C., & Tremblay, M.S. (2010). Advancing the debate on 'fitness testing' for children: Perhaps we're riding the wrong animal. *Pediatric Exercise Science*, 22(2), 176-182.

Naughton, G., Carlson, J.S., & Greene, D. (2006). A challenge to fitness testing in primary schools. *Journal of Science and Medicine in Sport*, 9(1), 40-45.

Read, J.C. (2008). Validating the fun toolkit: An instrument for measuring children's opinions of technology. *Cognition, Technology & Work*, 10(2), 119-128.

Read, J.C., MacFarlane, S., & Casey, C. (2002). Endurability, engagement and expectations: Measuring children's fun. *Interaction Design and Children*, 2, 1-23.

Risden, K., Hanna, E., & Kanerva, A. (1997). Dimensions of intrinsic motivation in children's favorite computer activities. *Society for Research in Child Development*, Washington, DC, 7.

Rowland, T.W. (2007). Fitness testing in the schools: Once more around the track. *Pediatric Exercise Science*, 19(2), 113.

Shephard, R.J., & Trudeau, F. (2010). The legacy of physical education: Influences on adult lifestyle. *Pes*, 12(1) *Pediatric Exercise Science*, 12(1), 34-50.

Silverman, S., Keating, X.D., & Phillips, S.R. (2008). A lasting impression: A pedagogical perspective on youth fitness testing. *Measurement in Physical Education and Exercise Science*, 12(3), 146-166.

Sim, G., MacFarlane, S., & Read, J. (2006). All work and no play: Measuring fun, usability, and learning in software for children. *Computers & Education*, 46(3), 235-248.

Stanec, A. D. S. (2009). The theory of planned behavior: Predicting teachers' intentions and behavior during fitness testing. *Journal of Teaching in Physical Education*, 28(3), 255-271.

Tomlinson, D., von Baeyer, C. L., Stinson, J. N., & Sung, L. (2010). A systematic review of faces scales for the self-report of pain intensity in children. *Pediatrics*, 126(5), e1168-98.

Vandorpe, B., Vandendriessche, J., Lefèvre, J., Pion, J., Vaeyens, R., Matthys, S. a., Lenoir, M. (2011). The Körperkoordinations Test für kinder: Reference values and suitability for 6-12-year-old children in flanders. *Scandinavian Journal of Medicine & Science in Sports*, 21(3), 378-388.

Webb, L., Quennerstedt, M., & Öhman, M. (2008). Healthy bodies: Construction of the body and health in physical education. *Sport, Education and Society*, 13(4), 353-372.

Whitehead, M. (2013). Definition of physical literacy and clarification of related issues. *ICSSPE Bulletin*, 65(1.2).

Wrotniak, B.H., Epstein, L.H., Dorn, J.M., Jones, K.E., & Kondilis, V.A. (2006). The relationship between motor proficiency and physical activity in children. *Pediatrics*, 118(6), e1758-65.

Zaman, B., Abeele, V. V., & De Grooff, D. (2013). Measuring product liking in preschool children: An evaluation of the smileyometer and this or that methods. *International Journal of Child-Computer Interaction*, 1(2), 61- 70.





# 6

## Validity and feasibility of an obstacle course to assess fundamental movement skills in a pre-school setting

Klingberg, B., Hoeboer, J. J. A. A. M., Schranz, N., Barnett, L. M., De Vries, S. I., & Ferrar, K. (2019). Validity and feasibility of an obstacle course to assess fundamental movement skills in a pre-school setting.

*Journal of Sports Sciences, 1-9.*

Failure to master age-appropriate fundamental movement skills (FMS) at a young age can limit motor skill competence affecting health. Assessments often have issues with feasibility and implementation in a field setting. As such, the purpose of this study was to investigate the validity and feasibility of the Athletic Skills Track (AST), in a pre-school setting. For the validation study sixty-five 3-6-year-old children (25 boys and 40 girls) from five pre-schools across Adelaide, Australia participated. Correlations and linear regression analysis (adjusted for age and gender) were used to investigate the association between the time to complete the AST and the raw score of the Test of Gross Motor Development 2 (TGMD-2). For the feasibility study pre-school staff completed a semi-structured interview regarding the feasibility of the AST. The AST took less than a minute per child and the TGMD-2 around 20 minutes for two children. There was a strong negative correlation ( $r = -0.63$ ,  $P < 0.01$ ) between the AST scores and the TGMD-2 scores. All five staff reported strengths of the AST to be its short administration time, setup and appropriateness. These results suggest that the AST could be a feasible and valid method of FMS assessment in Australian pre-schools.

## 6 Validity and feasibility of an obstacle course to assess fundamental movement skills in a pre-school setting

### 6.1 Introduction

Fundamental movement skills (FMS) are typically categorised as locomotion (e.g. running, jumping), object-control (e.g. throwing, catching) and stability skills (e.g., balance and body rolling) (Gallahue, Ozmun, & Goodway, 2012). According to Clark and Metcalfe (2002), FMS are considered to be a key component of motor competency (MC).

The importance of FMS and MC in children has been repeatedly studied worldwide. From this continual investigation, three important relationships appear; better FMS/MC are related to greater cardiorespiratory and musculoskeletal fitness (Barnett, Van Beurden, Morgan, Brooks, & Beard, 2008; Lubans, Morgan, Cliff, Barnett, & Okely, 2010), healthier weight status in children (Barnett et al., 2016; Cattuzzo et al., 2014; D'Hondt et al., 2014) and finally, better FMS and MC are associated with higher physical activity (PA) levels (Cohen, Morgan, Plotnikoff, Callister, & Lubans, 2015; Lloyd, Saunders, Bremer, & Tremblay, 2014; Logan, Kipling Webster, Getchell, Pfeiffer, & Robinson, 2015).

The critical time period for the development of FMS is early childhood (Clark, 1994; Seefeldt, 1980) with recent systematic reviews concluding that FMS interventions in early childhood are an effective way to improve FMS (Engel, Broderick, van Doorn, Hardy, & Parmenter, 2018; Logan, Robinson, Wilson, & Lucas, 2012). Therefore, it is important to screen and monitor FMS levels of pre-school aged children (3-6 years).

There are two main approaches used to assess FMS, process-oriented and product-oriented. Process-oriented assessments consider how a movement is performed giving a qualitative description, whereas product-oriented assessments consider quantifiable outcomes (Gabbard, 2012). For example, a product-oriented assessment might record the time taken for a child to run 10m whereas a process-oriented assessment would examine how a child ran 10m (by observing the quality of the leg and arm movements) (Logan, Barnett, Goodway, & Stodden, 2016).

Recent studies found significant positive (but only moderate) correlations (0.37-0.49) between process- and product-oriented assessments measuring the same construct (Logan, et al., 2016; Logan, Robinson, & Getchell, 2011; Valentini, 2015). Some researchers state that both types of assessments (product- and process-oriented) should be conducted to provide a more complete evaluation of motor skill competence (Logan, et al., 2016), however possibly counterproductive to being a feasible assessment for use.

There are numerous FMS assessments available for use in a clinical setting to identify disabilities or delays, such as the Körperkoordinationstest für Kinder (KTK), Motor Assessment Battery for Children-2 (MABC-2), Motor-Proficiency-Test (MOT 4-6), and the Test for Gross Motor Development-2 (TGMD-2). However, none have been recognised as being the most reliable and valid for use in a practical setting such as physical education classes or pre-school (Cools, Martelaer, Samaey, & Andries, 2009; Piek, Hands, & Licari, 2012). These assessments differ extensively in factors that can affect the feasibility of the use of FMS assessments in pre-school settings, such as administration time, training, space and equipment requirements (Klingberg, et al., 2018). More feasible assessments could allow for evaluations to be conducted by non-clinicians and hence may result in earlier interventions for children with poor FMS.

The Athletic Skills Track (AST) is a recently developed product-oriented obstacle course MC assessment that is able to address some of these issues (Hoeboer, Krijger, Savelsbergh, & De Vries, 2017). The AST is designed to measure 25 to 30 children in a one-hour PE lesson. The AST is fundamentally different to most conventional MC tests because it doesn't consist of isolated movements but rather, the obstacle course is designed around a series of 5–7 concatenated FMS (e.g., hopping and balancing). The underlying structure of the AST is based upon the idea that coordinative abilities (e.g., coupling, spatial orientation, and balance ability) are crucial in developing MC (Wormhoudt, Teunissen, & Savelsbergh, 2012) and are one of the underlying components of FMS (Barnett et al., 2016). The AST has been criterion tested against the product-oriented Körperkoordination-Test für Kinder (KTK) in 930 (4- to 12-year old) Dutch children, with results in young children (4- to 6-year olds) on the AST-1, favourable ( $r = -0.75$ ) (Hoeboer et al., 2017). The test-retest reliability of the AST-1 was high (ICC = 0.88 (95% CI: 0.78-0.93) and the internal consistency above the acceptable level of Cronbach's  $\alpha > 0.70$  ( $\alpha = 0.76$ ).

Yet, the validity of the AST against a process-oriented assessment, such as the Test of Gross Motor Development 2 (TGMD-2), is unknown. It is also unclear if the AST is feasible in the pre-school setting. As such, the following study aims to assess: i) the validity of the AST when compared to a process-oriented assessment in a population of pre-school aged children (age 3-6 years old) and ii) the feasibility of the AST when administered by pre-school staff in a field-based setting.

## 6.2 Methods

This study consisted of two concurrently run studies. The first was a validation study, using a cross-sectional design in which associations were tested between scores from the AST (Hoeboer, Krijger, et al., 2017) and a reference standard, the TGMD-2 (Ulrich, 2000). Secondly, feasibility of the AST was investigated within a pre-school setting, using semi-structured interviews to gather feasibility data from pre-school staff and gaining feedback from the children on their enjoyment of the AST using a Smileyometer. Both studies were conducted in pre-schools within 2-hours' drive of metropolitan Adelaide, South Australia, Australia.

### 6.2.1 Validation study

#### 6.2.1.1 Participants

Sixty five pre-school children (25 = boys, 40 = girls) aged 3-6 years old; independently ambulant; and able to understand English were recruited from pre-schools selected randomly from each of the socio-economic quartiles (Australian Bureau of Statistics, 2011). A staff member was also nominated by the director of each pre-school site to assist with research administration.

Children were excluded if they had current injury; pain; medication that may affect ability to participate; and/or a medical history making it unsafe to participate. Using a power calculation, it was determined that a sample size of 65 children was required for 80% power with an alpha of 0.05 and an expected correlation in a single tailed test of at least a medium effect size (0.3) between the reference test (TGMD-2) and the AST. The recruitment process was ongoing from October 2016 until January 2017.

#### 6.2.1.2 Procedure

**Researcher training:** Before assessment administration the research team underwent training with the AST developer, Joris Hoeboer (The Hague University of Applied Sciences, The Netherlands) in October 2016 and was supplied with the Standard Operating Procedures Manual (translated into English for this study). This translation was checked by the developer during the training. This training session took approximately 30 minutes to complete. The developer of the AST also ensured the research team were able to consistently educate others regarding the AST due to experience in training in the Netherlands. The research team also received training in the administration of the TGMD-2 from a qualified paediatric physiotherapist.

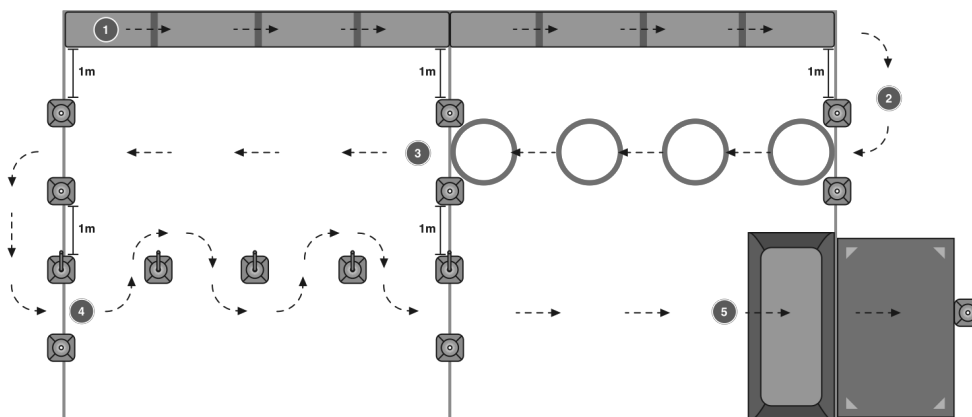
**Pre-school staff training:** The research team trained the nominated staff member from each pre-school in administering and setting up the AST at their pre-school site in a 30 minutes training session. The training sessions comprised of practicing set up, demonstration and scoring clarification. The training session also involved tuition in the appropriate levels of encouragement and amount and type of instruction to be provided, based on the Standard Operating Procedures Manual, and with experience from the training sessions completed in the Netherlands by the developer.

Data collection occurred between November 2016 and February 2017. On the day of data collection, the children were assessed using the AST and the TGMD-2. When not being tested the children were participating in regular activities during a pre-school day. The order in which the children performed each assessment was varied. The TGMD-2 was administered and filmed by the research team, taking approximately 20 minutes for two children to complete, the videos were scored later. The AST was set up, administered and timed by the staff member and supervised by the research team. The AST was scored live. The AST was completed in outdoor spaces available in the pre-school and the TGMD-2 was completed in indoor and outdoor spaces.

#### *6.2.1.3 Measurements*

##### **Athletic Skills Track**

The AST requires children to perform five to seven discrete locomotor and stability skills. These skills are set up in an obstacle course which is demonstrated by the assessor; see Figure 6.1 for course set up. Then the children complete two practice trials and one test attempt as fast as possible while still performing each skill correctly, taking less than a minute to complete the trial per child. There are three age-specific versions of the AST. In this study, the AST-1 developed for 4-to 6-year old children was administered, which consisted of five skills: balancing, forward leaps, alligator crawl (hands and feet), slaloming (weaving between poles) and clambering (climbing over a box). The score obtained was the time to complete the track (Hoeboer, et al., 2017). The time can be converted to a motor quotient (Hoeboer et al., 2018) however raw scores of the AST were used in the analysis due to the quotient being developed in a different country.



**Figure 6.1** AST-1 schematically displayed. (1) Walking, (2) Traveling jumps, (3) Alligator crawl, (4) Slaloming, and (5) Clambering

### Test of Gross Motor Development 2

The TGMD-2 is a process-oriented FMS assessment for children aged between 3-10 years old (Ulrich, 2000). The TGMD-2 consist of six locomotor skills: run, gallop, hop, leap, horizontal jump, and slide; and six object control skills: striking, dribble, catch, kick, overhand throw, and underhand roll. The assessor demonstrates these skills, then participants perform each skill twice after a practice trial, with the skills scored using set performance criteria. This assessment is reported to take 35 minutes to complete per child (Ulrich, 2000) (shorter in this study as performance video recorded and scored later). These scores are summed to form a raw score which can then be converted to a standard score or gross motor quotient (Ulrich, 2000), however raw scores of the TGMD-2 were used in the analysis as this is recommended for research purposes (Ulrich, 2000). In this study to ensure interrater accuracy, the coder (author removed for anonymity) and an experienced coder (author removed for anonymity) scored a random 10% of videos (seven children's videos of all 12 skills). Inter-rater reliability for the total raw score and subset scores (locomotor and object control) were assessed using intra-class correlation (ICC), using a two way mixed model for consistency for single measures (Koo and Li, 2016). The ICC was excellent for locomotor (ICC = 0.97, 95% CI = 0.85, 0.99), object control (ICC = 0.98, 95% CI = 0.91, 0.99), and total scores (ICC = 0.98, 95% CI = 0.92, 0.99).

## **Data analysis**

The time taken to complete the AST, in seconds accurate to a tenth of a second, was measured with a stopwatch and recorded onto a data sheet. The score for each criteria of the TGMD-2 was also recorded in raw score form onto the data sheet and used to generate the subset score and total raw score of the TGMD-2 for each participant. The data sheet was then transferred to SPSS version 21 for analysis.

Data analysis to detect normality was conducted by plotting histograms, examining the skewness and kurtosis, and using Shapiro-Wilks normality test. Due to normality assumptions being met, the strength of the correlation between the time to complete the AST and the raw scores of the TGMD-2 (total and subsets) was analysed using Pearson correlation coefficient. The data were checked against the six assumptions necessary for linear regression analysis, and as it met the assumptions, multiple linear regression was then calculated to predict the TGMD-2 score based on the AST score, with age (decimal age) and gender as covariates, as they are both known correlates of FMS (Barnett et al., 2016; Sterdt, Liersch, & Walter, 2014). A correlation of 0.20-0.39 was considered weak, 0.40-0.59 moderate, 0.60-0.79 strong and >0.80 very strong (Evans, 1996).

### 6.2.2 Feasibility study

#### 6.2.2.1 *Participants*

The sample group for the semi-structured interviews comprised one staff member from each pre-school who was nominated by the director and whom agreed to participate. The sample group for the Smileyometer (to determine child enjoyment) consisted of the children who completed the AST in the validation study. The Smileyometer was administered to 64 children, as one child declined due to behavioural issues.

#### 6.2.2.2 *Procedure*

After the pre-school staff had administered the AST, they were requested to complete a semi-structured interview that took approximately 20 minutes for each staff member. Five interviews took place in total, one with each staff member involved. This interview was audio recorded then transcribed and sent back to the staff member for confirmation and clarification as required. These confirmed transcripts were considered complete and used for data analysis. Immediately after completing the AST children were asked to individually score how much they enjoyed performing the AST on the Smileyometer (Read, MacFarlane, & Casey, 2002).

## Measurements

The semi-structured interview questions for staff members were created by the research team and framed around the AST and feasibility critical concepts (Bowen et al. 2009). These questions were open ended and guided to cover the topics of acceptability, demand, implementation, practicality, and integration; see Table 6.1 for interview questions.

**Table 6.1** Semi structured interview questions.

<b>Feasibility Concepts</b>	<b>Questions</b>
<b>Acceptability</b>	To what extent do you think this new assessment tool is suitable and attractive for use by staff members? How appropriate do you think it was for the children? How successful do you feel it was for the children? What did you find positive or negative about using this assessment? How best do you think the tool could be trained to staff members/ how was your training?
<b>Demand</b>	To what extent is the new assessment tool likely to be used by your pre-school? How much of a demand do you think there is for a tool like this?
<b>Implementation/practicality</b>	What are your thoughts regarding how easy or hard it would be for you and other staff to use the tool with the children? Would it be practical? How do you feel about the space/time/equipment required? Was there anything restricting its use in your pre-school?
<b>Integration</b>	What are your thoughts on how well the testing tool could be integrated into your current curriculum/plans? If the tool is shown to be accurate what benefits or negatives do you think there would be to the children?

The Smileyometer, that was used to assess children's enjoyment of the AST, is based around a 1-5 Likert scale using smiley faces to represent 'awful' to 'brilliant' (Read, MacFarlane, & Casey, 2002) (Figure 6.2). The usefulness, reliability and age effect of the Smileyometer have all been investigated and shown to be positive and recommended for continued use (Read, 2008). The Smileyometer has also been used in studies involving children (aged 7-8 years) rating enjoyment of different activities (Sim, MacFarlane, & Read, 2006) as well as to assess children's enjoyment of the AST in a sample of 4- to 12-year old Dutch children in a school setting (Hoeboer, De Vries, Mast, & Savelsbergh, 2017)



**Figure 6.2** Smileyometer scale.



### 6.2.2.3 Data Analysis

The transcribed interviews with the staff members were analysed by the research team for feasibility concepts, based on literature (Bowen et al., 2009), and described narratively. This was done by collating the transcribed interviews and comparing the responses from each staff member, looking for common answers or contradicting answers and/or key important feedback. The children's score on the visual Smileyometer was recorded as a numerical 1-5 score in the data sheet.

## 6.3 Results and Discussion

### 6.3.1 Validity results

#### 6.3.1.1 Participants

Parental consent was obtained from 74 children (consent rate of 51%), with data being collected from a total of 65 children, 25 boys and 40 girls (aged 3-6 years). The nine children who had parental consent from whom data was not collected declined to participate on the day of testing. The children were sourced from five pre-schools within 2-hours of metropolitan Adelaide, South Australia, with a pre-school response rate of 63%. See Table 6.2 for comprehensive descriptive data of results.

**Table 6.2** Descriptive data.

	<b>n</b>	<b>Mean</b>	<b>SD</b>	<b>Min.</b>	<b>Max.</b>
<b>Age</b> (decimal years)	65	4.66	0.54	3.5	5.5
<b>TGMD-2</b> (possible range 0-96)	65	41	11	8	65
<b>Locomotor</b> (possible range 0-48)	65	23	6	8	39
<b>Object control</b> (possible range 0-48)	65	18	5	0	29
<b>AST</b> (seconds)	65	27.5	6.0	17.3	40.9
<b>Smileyometer</b> (possible range 1-5)	65	4.5	0.8	2	5

**Relationship between AST and TGMD-2**

There was a strong negative correlation between the time to complete the AST and the total raw score from the TGMD-2 ( $r = -0.63, p < 0.01$ ). A moderate negative correlation between the AST and the locomotor subset ( $r = -0.52, p < 0.01$ ) and a strong negative correlation with the object control subset ( $r = -0.62, p < 0.01$ ) were also found.

The multiple linear regression, after adjustment for age ( $p < 0.038$ ) and gender (not significant), demonstrated the AST score was a statistically significant predictor of the TGMD-2 total raw score ( $p < 0.001, \beta = -0.55$ ). The total amount of adjusted variance explained in the model was 35% (adjusted  $R^2 = 0.35$ ). This process was repeated for the subsets of the TGMD-2. For the object control subset after adjustment for age and gender (both not significant), the AST score was shown to be a statistically significant predictor of the object control subset ( $p < 0.001, \beta = -0.53$ ) with the total amount of adjusted variance explained in the model was 32% (adjusted  $R^2 = 0.32$ ). For the locomotor subset after adjustment for age ( $p < 0.019$ ) and gender (not significant), the AST score was shown to be a statistically significant predictor of the locomotor subset ( $p < 0.001, \beta = -0.46$ ) with the total amount of adjusted variance explained in the model was 27% (adjusted  $R^2 = 0.27$ ). See Tables 6.3 through 6.8 for results from regression analysis.

**Table 6.3** Multiple Linear Regression model summary for TGMD-2 total score

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
	0.62	0.38	0.35	7.99

Predictors: (Constant), AST, Age, Male/Female.

**Table 6.4** Multiple Linear Regression coefficients for TGMD2 total score

Model	Unstandardized Coefficients		Standardized Coefficients	t	Sig.
	B	Std. Error	Beta		
(Constant)	50.30	9.83		5.12	.000
AST	-.94	.18	-.55	-5.15	.000
Male/Female	1.39	2.16	.07	.64	.523
Age	3.25	1.53	.22	2.12	.038

**Table 6.5** Multiple Linear Regression model summary for TGMD2 object control subset

<b>Model</b>	<b>R</b>	<b>R Square</b>	<b>Adjusted R Square</b>	<b>Std. Error of the Estimate</b>
1	0.59	0.35	0.32	4.12

Predictors: (Constant), AST, Age, Male/Female.

**Table 6.6** Multiple Linear Regression coefficients for TGMD2 object control subset

<b>Model</b>	<b>Unstandardized Coefficients</b>		<b>Standardized Coefficients</b>	<b>t</b>	<b>Sig.</b>
	<b>B</b>	<b>Std. Error</b>	<b>Beta</b>		
(Constant)	28.53	5.07		5.63	.000
Age	.79	.79	.11	.10	.324
Male/Female	-.91	1.11	-.09	-.82	.415
AST	-.45	.09	-.53	-4.83	.000

Dependent Variable: Object control subset.

**Table 6.7** Multiple Linear Regression model summary for TGMD2 locomotor subset

<b>Model</b>	<b>R</b>	<b>R Square</b>	<b>Adjusted R Square</b>	<b>Std. Error of the Estimate</b>
1	0.55	0.30	0.27	5.31

Predictors: (Constant), AST, Age, Male/Female.

**Table 6.8** Multiple Linear Regression coefficients for TGMD2 locomotor subset

<b>Model</b>	<b>Unstandardized Coefficients</b>		<b>Standardized Coefficients</b>	<b>t</b>	<b>Sig.</b>
	<b>B</b>	<b>Std. Error</b>	<b>Beta</b>		
(Constant)	21.77	6.54		3.33	.001
Age	2.46	1.02	.27	2.42	.019
Male/Female	2.30	1.44	.18	1.60	.115
AST	-.49	.12	-.46	-4.01	.000

Dependent Variable: Locomotor subset.

### 6.3.2 Validity discussion

The strength of relationship between assessments was encouraging considering the AST is still being validated and this is the first time it was compared against a process-oriented assessment. A study conducted by Hoeboer et al. (2017) with 930 Dutch children obtained similar results ( $r = -0.75$ ) when comparing the AST to a product-oriented assessment, the KTK. The reduced strength of relationship between the assessments in the current study may reflect the weaker association between a product- and process-oriented assessment and potentially, the smaller sample size. Interestingly, the correlations reported in this study are stronger than that found in previous comparisons of process- and product-oriented assessments, in particular the TGMD-2 and the MABC-2, with correlations ranging from 0.37-0.49 (Logan, et al., 2011; Valentini, 2015). This may suggest that the product-oriented AST is able to produce similar information regarding motor skill competence as the process-oriented TGMD-2. This might be the result of the underlying construct on which the product (AST) and process (in general) MC instruments are based on. The AST is developed based on the idea that coordinative abilities are crucial in the development of MC (Hoeboer, Krijger, et al., 2017). These coordinative abilities might also be measured indirectly with a process orientated assessment such as the TGMD-2.

When considering the subsets, there was a moderate relationship ( $r = -0.52$ ) with the locomotor subset and a strong relationship ( $r = -0.62$ ) with the object control subset of the TGMD-2 against the AST. These results are also reflected in the regression results after adjusting for age and gender. The stronger relationship with the object control subset is counterintuitive as there are no object control skills specifically assessed in the AST. There are a few possible explanations for this, overall, the scores for the object control subset were lower than the scores for the locomotor subset, reflective of the young age group where object control skills are only just developing (Clark, 1994). This may offer a possible explanation, in that the AST scores were more aligned with the lower scoring subset regardless of the skills within it. Alternatively, it could be that the object control skills are more complex than locomotor skills and thus relate better to an assessment that needs to be completed in a coordinative manner as fast and as proficiently as possible. Furthering on this, as the AST contains balance and stability skills, both of which are critical during the early childhood years (Cliff, Okely, Smith, & McKeen, 2009; Robinson, Wadsworth, & Peoples, 2012) this may explain the stronger relationship (Rudd et al., 2015). This aligns with motor development literature, as locomotor skills develop earlier than object control/ manipulative skills (Robinson et al., 2015). The relationship between the AST and FMS subsets has yet to be studied elsewhere for comparison. However,

a previous study looking at the MOT 4-6, containing locomotor and object control subsets, and the KTK (which assess motor coordination), can be considered. This study found that the relationship between the KTK and the locomotor subset ( $r = 0.56$ ) was stronger than with the object control subset ( $r = 0.37$ ) (Bardid et al., 2016) which is what might be expected as the KTK does not contain object control skills. The findings of this current study however appears to contrast the little literature available (Bardid, et al., 2016) and as such further investigation is required.

### 6.3.3 Feasibility results

#### 6.3.3.1 Participants

All five staff members (5 females, aged 32-34 years) involved in this study, four teachers (diploma level) and one early childhood worker (certificate level), completed the semi-structured interview.

#### **Staff feedback**

In response to questions regarding acceptability, all staff mentioned that the AST was quick and easy to use and was at the appropriate skill level for the children. Four staff commented that the style (obstacle course) was enjoyable for the children, three commented that the space requirements were appropriate, and one staff member mentioned that the skill slaloming (weaving between cones) may be too advanced for younger children. Three out of five staff commented they wanted more information regarding each skill assessed. Three staff members also mentioned that the map of the course could be made clearer to aid in setting up and understanding the course.

Responding to questions surrounding the demand for the AST, four staff commented that something like the AST is needed to allow staff to assess children. All staff noted that they would likely use the AST in their pre-school again. One staff member said they would be more likely to use the AST again if support material was provided (e.g., activity suggestions) to help the children improve their motor skills.

Asking about the implementation and practicality of the AST, all five staff thought the assessment would be accessible for all staff members to use, however one mentioned the physical requirements of demonstrating the courses may be inappropriate for some staff. Staff all noted that time to administer would make it practical and three thought the equipment was practical, with a fourth staff member commenting that their pre-school would purchase the equipment (foam climbing box) if necessary. Reported restrictions to the implementation of

the AST were site specific including, behaviour issues of children at specific sites, staff training, space requirements and difficulty of use with younger children. Two staff members mentioned that their pre-school site lacked the specific equipment for the AST (foam climbing box).

With regard to integration, staff commented that if in line with the goals of their pre-school, the AST could fit into their curriculum and would benefit the children.

### **Smileyometer**

The mean score of the children's rating of their enjoyment of the AST was 4.5 out of 5, with only one child scoring below a three.

#### 6.3.4 Feasibility discussion

The feasibility study investigated pre-school staff's opinions on the use of the AST in their pre-schools. Overall the results from the feasibility study suggest that staff found the AST to be feasible, therefore it could be considered an alternative to current FMS assessments such as the TGMD-2.

All staff felt the administration time for the AST was feasible. This is in line with a recent systematic review (reference removed for anonymity) where the AST scored the highest with regards to administration time compared to commonly available FMS assessments. Overall, the staff thought the skills were appropriate for the children; although one staff member raised concerns that the skills within the AST are slightly different to commonly taught skills in Australia. Since the AST was developed in Europe, different skills can be considered fundamental in different cultural settings, and hence differ worldwide (Barnett, Stodden, et al., 2016). In fact, a recent conceptual model suggests that culture and geographical location are important to consider in movement skill development (Hulteen, Morgan, Barnett, Stodden, & Lubans, 2018). The implication being that staff may have to decide whether there is value in assessing skills not typical for pre-school teaching and activities in Australia such as slaloming. However, slaloming represents movement coordination (lateral weight shifting and moving in different directions) required in all ball games worldwide and is similar to a dodge, which is assessed in an Australian FMS assessment (New South Wales Department of Education and Training, 2000). So, in this respect it could still be considered as a relevant skill to assess in an Australian context.

The majority of staff reported wanting more information regarding what skills they are assessing and how to promote those skills. This is not an entirely new concept as the

incorporation of possible intervention strategies appears in previous assessments including the MABC-2, PDMS-2 and the MOT 4-6 (Cools, et al., 2009). It is predicted that demand for interventions coupled with assessments may increase as more pre-schools implement their own self-administered FMS assessments (Kambas and Venetsanou, 2014; Lam, Ip, Lui, & Koong, 2003).

Regarding the implementation of the AST, a concern was highlighted with regards to the physical demand on staff members to demonstrate the course. This may be a consideration affecting how well this assessment can be implemented across all pre-schools. This could be overcome using iPads with expert demonstrations to show children the course, as done previously for an FMS obstacle course for 10-14-year-old children (Tyler, Foweather, Mackintosh, & Stratton, 2018). Although the efficacy of this method in terms of pre-school children understanding how to mimic the skill in real space and time would need to be tested. These suggestions will be provided to the developer of the AST for their consideration for implementation in future research.

#### 6.3.5 Strengths and limitations

A strength to the methods of this study is it clear reporting on the feasibility of a new FMS assessment, this was identified to be usually reported poorly by studies (Klingberg, et al., 2018). We reported clearly on the training we provided to the teachers involved in the study which is critically important but also rarely reported (Lander, Eather, Morgan, Salmon, & Barnett, 2017).

The obstacle course design of the AST appears to have two possible sides. On one side, the use of an obstacle course makes this assessment closer to real world FMS which is though not to be captured in common FMS assessments (Tidén, Lundqvist, & Nyberg, 2015), possibly improving its ecological validity. However, there is also the possibility that being an obstacle course design means children may practice it using it for training rather than testing. Maybe improving their ability to complete the track rather than actually improving their FMS, i.e. habituation. Research does however support the use of obstacle courses to improve FMS skills (Lander, Morgan, Salmon, & Barnett, 2017). In fact, the recent intervention by Lander et al used an obstacle course in this way, with evidence of successful improvement in FMS and perception of FMS. Further study will need to be done to investigate these theories but initial results from Hoeboer et al. (2017) showed high degree of test re-test reliability (ICC = 0.881, 95% CI: 0.780–0.934) when children completed the test twice two weeks apart.

There is currently no normative data for the AST to create cut off ranks to allow interpretation of the children's scores in the AST. Past studies of the AST in the Netherlands has shown that the AST scores could categorise and discriminate children into seriously motor disordered, moderately motor disordered and normal motor gifted groups as indicated by the MQ of the KTK (Hoeboer et al., 2016). Also, specific to this study we had one child score well below the other in the TGMD-2, 15 points lower than next closest score scoring 0 for object control subset. This child was anecdotally also identified by the AST as they scored 4.7 seconds slower than the next closest child, with all other scores split by less than a second. Future research may wish to determine the discriminant properties of the AST at identifying FMS levels to allow for targeted intervention by the user.

In the validation study the recruitment process was sufficient to reach power, but insufficient to account for potential sample and location bias. Due to a lack of funding, pre-schools were only considered if they were within a 2-hour drive of metropolitan Adelaide. This means the data may not represent all South Australia, particularly the outer regional areas.

Also, even though the validation test was carried out in typical pre-school settings, the variation in each setting may have influenced the between pre-school results. Each pre-school had different areas available for the filming of the TGMD-2 and setup of the AST (outside on grass, paving, bark chips). Future, larger validation studies will have the potential to account for differences at the pre-school level.

Finally, as the testing was occurring throughout the day, the time of testing may have affected their energy levels. So, if they were assessed at the end of the day they may have performed poorer relative to others due to low energy levels.

#### 6.4 Conclusion

This study has addressed an important topical area regarding the validity and feasibility of the AST in the pre-school setting. The results from this study suggest that the AST could be a feasible and valid FMS assessment to use in Australian pre-schools. With the development of normative data and cut off levels for pre-school aged children in Australia this assessment could provide pre-school staff the ability to formally assess the FMS of their children without referring to clinical experts, thus allowing for more frequent evaluations and earlier interventions for children with poor FMS.



## 6.5 References

Australian Bureau of Statistics. (2011). *Census of Population and Housing: Socio-Economic Indexes for Areas (SEIFA)*. (2033.0.55.001). Canberra: Australian Bureau of Statistics.

Bardid, F., Huyben, F., Deconinck, F. J., De Martelaer, K., Seghers, J., & Lenoir, M. (2016). Convergent and Divergent Validity Between the KTK and MOT 4-6 Motor Tests in Early Childhood. [Research Support, Non-U.S. Government].

Barnett, L. M., Lai, S. K., Veldman, S. L., Hardy, L., Cliff, D. P., Morgan, P. J., . . . Ridgers, N. D. (2016). Correlates of gross motor competence in children and adolescents: a systematic review and meta-analysis. *Sports Medicine*, 46(11), 1663-1688.

Barnett, L. M., Stodden, D. F., Cohen, K. E., Smith, J. J., Lubans, D. R., Lenoir, M., Dudley, D. (2016). Fundamental movement skills: An important focus. *Journal of Teaching in Physical Education*, 35(3), 219-225.

Barnett, L. M., Van Beurden, E., Morgan, P. J., Brooks, L. O., & Beard, J. R. (2008). Does childhood motor skill proficiency predict adolescent fitness? *Medicine & Science in Sports & Exercise*, 40(12), 2137-2144.

Bowen, D. J., Kreuter, M., Spring, B., Cofta-Woerpel, L., Linnan, L., Weiner, D., Fernandez, M. (2009). How we design feasibility studies. [Research Support, N.I.H., Extramural]. *American Journal of Preventive Medicine*, 36(5), 452-457.

Cattuzzo, M. T., dos Santos Henrique, R., Ré, A. H. N., de Oliveira, I. S., Melo, B. M., de Sousa Moura, M., Stodden, D. F. (2014). Motor competence and health related physical fitness in youth: A systematic review. *Journal of Science and Medicine in Sport*, 19(2), 123-129.

Clark, J. (1994). *Motor development Encyclopedia of Human Behavior* (3rd ed., Vol. 3rd edn). San Diego: Academic Press.

Clark, J., & Metcalfe, J. S. (2002). The mountain of motor development: A metaphor. *Motor development: Research and Reviews*, 2(163-190).

Cliff, D. P., Okely, A. D., Smith, L. M., & McKeen, K. (2009). Relationships between fundamental movement skills and objectively measured physical activity in preschool children. *Pediatric Exercise Science*, 21(4), 436-449.

Cohen, K. E., Morgan, P. J., Plotnikoff, R. C., Callister, R., & Lubans, D. R. (2015). Physical activity and skills intervention: SCORES cluster randomized controlled trial. *Medicine and Science in Sports and Exercise*, 47(4), 765-774.

Cools, W., Martelaer, K. D., Samaey, C., & Andries, C. (2009). Movement skill assessment of typically developing preschool children: a review of seven movement skill assessment tools. [Review]. *Journal of Sports Science & Medicine*, 8(2), 154-168.

D'Hondt, E., Deforche, B., Gentier, I., Verstuyf, J., Vaeyens, R., De Bourdeaudhuij, I., & Lenoir, M. (2014). A longitudinal study of gross motor coordination and weight status in children. *Obesity*, 22(6), 1505-1511.

Engel, A. C., Broderick, C. R., van Doorn, N., Hardy, L. L., & Parmenter, B. J. (2018). Exploring the Relationship Between Fundamental Motor Skill Interventions and Physical Activity Levels in Children: A Systematic Review and Meta-analysis. *Sports Medicine*, 1-13.

Evans, J. D. (1996). *Straightforward statistics for the behavioral sciences California*: Brooks/Cole Publishing.

Gabbard, C. (2012). *Lifelong motor development (6th ed.)* San Francisco: Pearson Benjamin Cummings.

Hoeboer, J., De Vries, S., Krijger-Hombergen, M., Wormhoudt, R., Drent, A., Krabben, K., & Savelsbergh, G. (2016). Validity of an Athletic Skills Track among 6-to 12-year-old children. *Journal of Sports Sciences*, 34(21), 2095-2105.

Hoeboer, J., De Vries, S., Mast, D., & Savelsbergh, G. (2017). Children's enjoyment of a motor skill test in physical education. *Journal of Physical Education Research*, 4(4).

Hoeboer, J., Krijger, M., Savelsbergh, G., & de Vries, S. (2017). Reliability and validity of an athletic skills track to assess motor competence among 4–12-year-old children. *Journal of Science and Medicine in Sport*, 20, 76-77.

Hoeboer, J. J. A. A. M., Ongena, G., Krijger-Hombergen, M., Stolk, E., Savelsbergh, G. J. P., & de Vries, S. I. (2018). The Athletic Skills Track: Age-and gender-related normative values of a motor skills test for 4-to 12-year-old children. *Journal of science and medicine in sport*, 21(9), 975-979.

Hulteen, R. M., Morgan, P. J., Barnett, L. M., Stodden, D. F., & Lubans, D. R. (2018). Development of Foundational Movement Skills: A Conceptual Model for Physical Activity Across the Lifespan. *Sports Medicine*, 1-8.

Kambas, A., & Venetsanou, F. (2014). The Democritos Movement Screening Tool for Preschool Children (DEMOST-PRE©): development and factorial validity. [Validation Studies]. *Research in Developmental Disabilities*, 35(7), 1528-1533.

Klingberg, B., Schranz, N., Barnett, L. M., Booth, V., & Ferrar, K. (2018). The feasibility of fundamental movement skill assessments for pre-school aged children. *Journal of Sports Sciences*, 1-9.

Koo, T. K., & Li, M. Y. (2016). A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *Journal of Chiropractic Medicine*, 15(2), 155-163.

Lam, M. Y., Ip, M. H., Lui, P. K., & Koong, M. K. (2003). How Teachers Can Assess Kindergarten Children's Motor Performance in Hong Kong. *Early Child Development and Care*, 173(1), 109-118.

Lander, N., Eather, N., Morgan, P. J., Salmon, J., & Barnett, L. M. (2017). Characteristics of teacher training in school-based physical education interventions to improve fundamental movement skills and/or physical activity: A systematic review. *Sports Medicine*, 47(1), 135-161.

Lander, N., Morgan, P. J., Salmon, J., & Barnett, L. M. (2017). Improving early adolescent girls' motor skill: A cluster randomized controlled trial. *Medicine & Science in Sports & Exercise*, 49(12), 2498-2505.

Lloyd, M., Saunders, T. J., Bremer, E., & Tremblay, M. S. (2014). Long-term importance of fundamental motor skills: a 20-year follow-up study. *Adapted Physical Activity Quarterly*, 31(1), 67-78.

Logan, S., Barnett, L. M., Goodway, J. D., & Stodden, D. F. (2016). Comparison of performance on process-and product-oriented assessments of fundamental motor skills across childhood. *Journal of Sports Sciences*, 35(7), 634-641.

Logan, S., Kipling Webster, E., Getchell, N., Pfeiffer, K. A., & Robinson, L. (2015). Relationship between fundamental motor skill competence and physical activity during childhood and adolescence: a systematic review. *Kinesiology Review*, 4(4), 416-426.

Logan, S., Robinson, L., & Getchell, N. (2011). The comparison of performances of preschool children on two motor assessments. *Perceptual & Motor Skills*, 113(3), 715-723.

Logan, S., Robinson, L., Wilson, A. E., & Lucas, W. A. (2012). Getting the fundamentals of movement: A meta-analysis of the effectiveness of motor skill interventions in children. [Review]. *Child: Care, Health and Development*, 38(3), 305-315.

Lubans, D. R., Morgan, P. J., Cliff, D. P., Barnett, L. M., & Okely, A. D. (2010). Fundamental movement skills in children and adolescents: review of associated health benefits. [Review]. *Sports Medicine*, 40(12), 1019-1035.

Piek, J. P., Hands, B., & Licari, M. K. (2012). Assessment of motor functioning in the preschool period. [Review]. *Neuropsychology Review*, 22(4), 402-413.

Read, J. (2008). Validating the Fun Toolkit: an instrument for measuring children's opinions of technology. *Cognition, Technology & Work*, 10(2), 119-128.

Read, J., MacFarlane, S., & Casey, C. (2002) Endurability, engagement and expectations: Measuring children's fun. Paper presented at the Interaction design and Children Conference.

Robinson, L. E., Stodden, D. F., Barnett, L. M., Lopes, V. P., Logan, S. W., Rodrigues, L. P., & D'Hondt, E. (2015). Motor competence and its effect on positive developmental trajectories of health. *Sports Medicine*, 45(9), 1273-1284.

Robinson, L. E., Wadsworth, D. D., & Peoples, C. M. (2012). Correlates of school-day physical activity in preschool students. *Research Quarterly for Exercise and Sport*, 83(1), 20-26.

Rudd, J. R., Barnett, L. M., Butson, M. L., Farrow, D., Berry, J., & Polman, R. C. (2015). Fundamental movement skills are more than run, throw and catch: the role of stability skills. *PloS one*, 10(10), p e0140224.

Seefeldt, V. (1980) Developmental motor patterns: Implications for elementary school physical education. Paper presented at the Refereed proceedings of Psychology of motor behavior and sport 1980, Illinois.

Sim, G., MacFarlane, S., & Read, J. (2006). All work and no play: Measuring fun, usability, and learning in software for children. *Computers & Education*, 46(3), 235-248.

Sterdt, E., Liersch, S., & Walter, U. (2014). Correlates of physical activity of children and adolescents: A systematic review of reviews. *Health Education Journal*, 73(1), 72-89.

Tidén, A., Lundqvist, C., & Nyberg, M. (2015). Development and initial validation of the NyTid test: A movement assessment tool for compulsory school pupils. *Measurement in Physical Education and Exercise Science*, 19(1), 34-43.

Tyler, R., Fowweather, L., Mackintosh, K. A., & Stratton, G. (2018). A Dynamic Assessment of Children's Physical Competence: The Dragon Challenge. *Medicine and science in sports and exercise*.

Ulrich, D. A. (2000). Test of gross motor development-2 Texas: Prod-Ed.

Valentini, N. (2015). Exploring Associations Between Motor Skill Assessments in Children With, Without, and At-Risk for Developmental Coordination Disorder. *Journal of Motor Learning and Development*, 3(1), 39-52.







# 7

## Reproducibility, validity and feasibility of motor competence assessment instruments: a systematic review

Hoeboer, J. J. A. A. M., Opstoel, K., Martelaer, K., Savelsbergh, G., & De Vries, S. I.

*(Under review)*

Introduction: Motor Competence (MC) is one of the key components in developing an active lifestyle. To identify children at risk and to define intervention goals, validated and reliable tools to measure MC are crucial. Therefore, the purpose of the present study is to systematically review the reliability, validity, and feasibility of MC measurement tools for children aged 2-to 18-years old. Method: A systematic search of the literature was performed in February 2018 in four databases resulting in 4,299 hits. After screening and backwards tracking 38 studies were included. The included studies were evaluated using a 10-item checklist. Results: There is strongest evidence for the reliability and validity of the M-ABC (n=16 studies) and the TGMD-2 (n=10). The M-ABC scores moderate to good on all items of reliability and validity. With respect to the TGMD-2 there is strong evidence for a good internal consistency, test-retest reliability, construct validity and a moderate inter-rater reliability. However, there is strong evidence for a poor concurrent validity. With respect to the AST there is modest evidence for a good internal consistency, test-retest reliability and concurrent validity. For the other eight reviewed measurement tools there is limited evidence for the clinimetric outcomes. Conclusion: In conclusion, there is quite some evidence for a good reliability and validity of several MC measurement tools for 2- to 18-years old children. However, the level of evidence differs. It depends on the target group, the purpose and the context which MC measurement tool can best be used.

## 7 Reproducibility, validity and feasibility of motor competence assessment instruments: a systematic review

### 7.1 Introduction

The decline of motor competence of children (MC) is a challenge because there is emerging evidence that supports the association between MC and all kinds of health-related outcomes (Robinson et al., 2015), such as physical activity (PA), cardiorespiratory fitness and weight status (Barnett, Van Beurden, Morgan, Brooks, & Beard, 2009; Lubans, Morgan, Cliff, Barnett, & Okely, 2010). MC is one of the key components in motor development (Lubans et al., 2010) and can be identified as proficiency in the range of fundamental movement skills (FMS) (e.g. running, catching, balancing) that should be mastered during the pre-school and primary school period (Branta, Haubenstricker, & Seefeldt, 1984; Gallahue, Ozmun, & Goodway, 2006; Gallahue & Donnelly, 2007). Those FMS are the foundation on which children develop more specialized, sport-specific tasks (e.g. chest-pass in basketball, serving in tennis) (Clark & Metcalfe, 2002). Mastering FMS during the childhood period makes lifetime physical activity possible (Hulteen et al., 2015).

The (pre-)school setting, especially the Physical Education (PE) lessons, present an important setting for developing MC, because all children participate in the PE lessons on a weekly basis. This setting provides personnel and resources, such as qualified teachers, equipment, space, and the ability to provide intervention programs to influence MC in a positive manner (Fairclough, Stratton, & Baldwin, 2002; Green, 2002; Kirk, 2005). In order to identify children at risk and to define intervention goals, reliable, valid, and feasible tools to measure MC are crucial (Okan, Pinheiro, Zamora, & Bauer, 2015).

Cools et al. reviewed seven MC measurement tools in 2009 (Cools, Martelaer, Samaey, & Andries, 2009). They concluded that the primary goal of the reviewed instruments is to detect deficiencies in motor skill development. Most of the studies using these tools hardly discuss the variation in motor skill development of typically developing children (Cools et al., 2009). Furthermore, Cools et al. (2009) did not execute a systematic review and therefore did not objectively selected the studies. Next to the selection procedure, the quality of the included studies was not taken into account when extracting and interpreting the data. Since the review of Cools et al. (2009) some new MC measurement tools have been introduced. Therefore, the purpose of the present study was to systematically review the reliability, validity, and feasibility of measurement tools to assess MC of normally developing children (aged 2-to 18-years old).

## 7.2 Method

A systematic review of the literature was conducted on measurement tools of motor competence among children from 2-to 18-year-old. To ensure the completeness of the review and to improve the reporting of this review, the checklist of the Preferred Reporting Items for Systematic reviews and Meta-Analyses (PRISMA statement) was used. The PRISMA flow diagram illustrates the inclusion and exclusion process (see Figure 7.1) (Moher, Liberati, Tetzlaff, & Altman, 2010).

### 7.2.1 Data sources and search strategy

Three researchers (J.H., K.O., S.d.V.) developed the search strategy. The searches were executed by an employee of the library of the The Hague University of Applied Sciences. Two researchers (J.H., K.O.) checked the initial results of the search and discussed the search outcomes. The search was performed in four databases: PubMed, Educational Resources Information Centre (ERIC), Academic Search Complete, and SPORTDiscus. The four databases were initially searched in February 2018 using a composite search term that comprised a combination of three terms for studies published between January 2000 and February 2018 to connect overlapping with the previously conducted review of Cools et al. (2009). To combine search terms Boolean operators (AND/OR) were used. The first search term was “fundamental movement skills”, the second term: “child” addressed the target population, and the third term focused on either one of the clinimetric properties of measurement tools. The following algorithm was used in PubMed:

*“(Fundamental Movement skills OR Motor Skill Competence OR Fundamental Motor Skills OR Motor Coordination OR Motor Competence OR Motor Skills OR Motor Ability OR Motor performance OR Motor proficiency) AND (children OR adolescence OR childhood OR adolescent OR child OR adolescent OR kids OR minor OR teenager OR youngster OR youth) AND (assessment OR test OR measurement) AND (validity OR reliability OR feasibility)”*

A similar search strategy was used in the other three databases. Additionally, reference lists of included full-text studies were searched manually (J.H.) for potentially additional relevant studies.

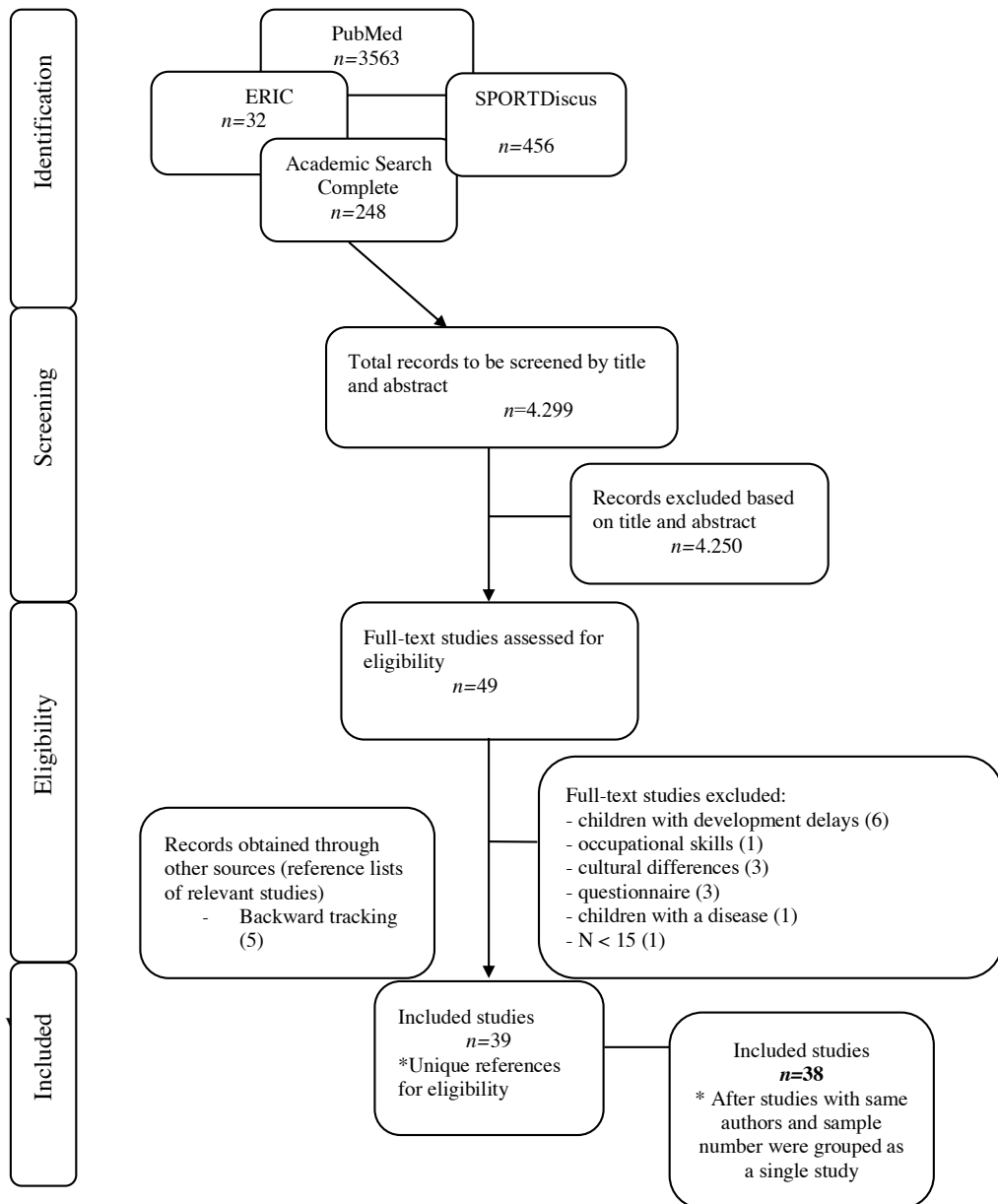


Figure 7.1 PRISMA flow diagram

### 7.2.2 Study selection

The following inclusion criteria were used to select relevant studies (see Table 7.1): (a) published between January 2000 and February 2018; (b) English language; (c) original publication in peer reviewed journal; (d) study reporting on reliability, validity, or feasibility of tools measuring motor competence; (e) normally developing children aged between 2 and 18 years old; (f) any setting; (g) any country; (h) motor skill assessments (physical test); (i) size of the study sample  $n \geq 15$  (based on COTAN (De Vries et al., 2009; Evers, Sijtsma, Lucassen, & Meijer, 2010)).

**Table 7.1** Inclusion and exclusion criteria.

<b>Criterion</b>	<b>Inclusion</b>	<b>Exclusion</b>
<b>Time</b>	January 2000 – February 2018	Studies before 2000 and after February 2018
<b>Language</b>	English	Any other language
<b>Type of publication</b>	Original papers published in peer reviewed journals	Any non-original publication, books, theses, editorials, letters to editors
<b>Focus of study</b>	Any study reporting on validity and reliability of instruments measuring motor skill competence	Any study that does not target on specific validity, reliability of instruments measuring motor skill competence
<b>Study population</b>	Normally developing Children $\geq 2$ and $\leq 18$ years of age	Children $< 2$ and $> 18$ years of age diseases and developmental delays
<b>Setting</b>	Any setting	Nil
<b>Country</b>	Any country	Nil
<b>Kind of measurement tool</b>	Motor skill assessments	Questionnaires, Survey, Perceived motor competence by children, parents or professionals
<b>Study size</b>	$N \geq 15$	$N < 15$

The search in the four databases resulted in 4.299 unique hits ( $n = 32$  in ERIC;  $n = 3563$  in PubMed;  $n = 248$  in Academic Search Complete,  $n = 456$  in SPORTDiscus). First, the title and abstract of these studies were screened independently by three authors (J.H., K.O., S.d.V.). This resulted in the exclusion of 4.250 studies. Next, 49 full-texts were analysed independently by two authors (J.H., K.O.). This resulted in the exclusion of another 15 studies (see Figure 7.1). Backward tracking resulted in five additional studies. Thus, this systematic review reports the evidence on the reliability, validity, and feasibility of MC measurement tools for 2- to 18-year old children based on 38 studies.

### 7.2.3 Data extraction and evaluation

All data of the included studies were extracted and evaluated independently by two authors (J.H., K.O.) using a 10-item checklist. The checklist is based on the COTAN criteria that describes a review process for evaluating the quality of testing (De Vries et al., 2009; Evers et al., 2010).

The checklist contains four items on study design, three items on reliability and three items on validity (see also Appendix 4). A clear description of the study design includes: characteristics of the study sample, protocol information, content of the measurement tool, and statistical analyses. All quality-related items were scored with 0, 0.5, or 1 point and summed per study (see Appendix 4). The reliability of a measurement tool concerns three concepts: internal consistency, intra-rater/ test-retest reliability, and inter-rater reliability. The validity was determined on three aspects: construct, content and concurrent validity. All the reliability and validity items were scored 0 point if the item was not described in the study and 1 point if the item was described in the study. The minimum score that a study could get on the 10-item checklist was 1 point, the maximum score was 10 points. Next, outcome measures of reliability and validity were rated as: poor (-), moderate (+/-), or good (+). The internal consistency was rated as good (+) if Cronbach's Alfa was higher than or equal to 0.70, moderate (+/-) if Cronbach's Alfa was between 0.60 - 0.70, and poor (-) if Cronbach's Alfa < 0.60 (Cronbach, 1951). Calculation of the intraclass correlation coefficient (ICC) was considered an adequate method to quantify intra- or inter-rater reliability. An ICC > 0.80 was considered to reflect good reliability (+), an ICC between 0.50-0.80 as moderate reliability (+/-), and an ICC < 0.50 as poor reliability (-) (Altman, 1991). Calculation of the correlation coefficient (r) was considered an adequate method to quantify construct or concurrent validity. A correlation coefficient of  $r > 0.70$  was considered as a good score (+),  $r$  between 0.50 - 0.70 was considered a moderate score (+/-), and  $r < 0.50$  as a poor score (-) (Hinkle, Wiersma, & Jurs, 2003). The two authors (J.H., K.O.) agreed on all ratings after discussion.

Next, the outcomes of the 38 studies were weighted for the quality of the study and summarized per measurement tool. Studies of poor quality provide less evidence for the reported outcomes than studies of good quality. In appendix 5 studies were given one character based on the outcome of the study (character: “-“= poor reliability or validity, “+/-” = moderate or mixed reliability or validity, “+” = good reliability or validity). The quality of the study is expressed in a second character (< 4 points = limited evidence, = no second character; 4 – 6 points = modest evidence, second character; > 6 points = strong evidence, third character). A third

character was also added if there are five or more studies that investigated one or more clinimetric properties of the MC measurement tool.

Besides data on the study design, reliability and validity of the MC measurement tool, data on the feasibility of the MC tool was also extracted from the studies. In this review five elements of feasibility were included in the review process based on a previously performed review (Cools et al., 2009): the assessor, test duration, equipment needed, cost of the measurement tool, and space needed to execute the measurement. With respect to the assessor a distinction was made between more clinical professionals (for example: clinical therapist, occupational therapist, paediatricians, physiotherapist) and school professionals, such as PE teachers, educational staff and preschool professionals.

### 7.3 Results

In total, 38 studies were included in this review (see Appendix 5). The selected studies describe aspects of reliability, validity, and feasibility of eleven MC measurement tools i.e.: Athletic Skills Track (AST):  $n = 2$  (Hoeboer, Krijger, Savelsbergh, & de Vries, 2017; Hoeboer et al., 2016), Bruininks-Oseretsky Test-2nd edition (BOT-2):  $n = 3$  (Bruininks, 1978; Bruininks, 2005), Canadian Agility and Movement Skill Assessment (CAMSA):  $n = 1$  (Longmuir et al., 2015), Democritos Movement Screening Tool for preschool children (DEMOST-pre):  $n = 1$  (Kambas & Venetsanou, 2014), Fundamental Movement Skill polygon (FMS-polygon):  $n = 1$  (Zuvela, Bozanic, & Miletic, 2011), Körperkoordinationstest für Kinder (KTK):  $n=3$  (Kiphard & Shilling, 2007; Kiphard & Schiling, 1974), Movement Assessment Battery for Children-2nd edition (M-ABC-2):  $n = 16$  (Henderson, 1992; Henderson, Sugden, & Barnett, 2007), Motoriktest für vier- bis sechsjährige Kinder (MOT 4-6):  $n = 2$  (Zimmer & Volkamer, 1987), Matorische BAsisKompetenzen in third-grade children (MOBAK):  $n = 2$  (Herrmann et al., 2015), Test of Gross Motor Development-2nd edition (TGMD-2):  $n = 10$  (Ulrich, 2000; Ulrich, 1985) and Zurich Neuromotor Assessment in Preschool Children (ZNA):  $n = 2$  (Largo et al., 2001).

#### 7.3.1 General characteristics and feasibility

The general characteristics of the MC measurement tools are presented in Table 7.2. The test goals of the included measurement tools are diverse but can be divided in two groups. In general, the AST, CAMSA, DEMOST-pre, FMS-polygon, MOBAK, and ZNA focus on measuring motor proficiency in typically developing children. The BOT-2, KTK, M-ABC-2,

MOT 4-6, and TGMD-2 aim to identify individuals with mild to severe motor coordination problems or to identify children that are behind.

Eight of the eleven measurement tools are product-orientated tools (AST, BOT-2, DEMOST-pre, FMS-polygon, KTK, M-ABC-2, MOBAK, MOT 4-6). Product-orientated assessments (Schulz, Henderson, Sugden, & Barnett, 2011) evaluate the outcome of a movement and have a quantified outcome, for example the time it takes to complete a certain task, or the number of successful attempts (Logan, Barnett, Goodway, & Stodden, 2017). The TGMD-2 is the only measurement tool that is process- orientated, the CAMSA and the ZNA are partly process- and partly product-orientated. Process-oriented assessments evaluate how a movement is performed, with qualitative characteristics that describe successful movement patterns (Logan et al., 2017).

With respect to the test settings and the assessor, five measurement tools are designed for a more clinical setting to be used by a psychologist, physical therapist, physiotherapist or paediatric (BOT-2, KTK, M-ABC-2, TGMD-2, ZNA). The other measurement tools are designed to be used in the school setting by educational staff, preschool professionals, sport coaches or PE teachers (AST, CAMSA, DEMOST-pre, FMS-polygon, MOBAK, MOT 4-6). In line with this, five measurement tools can be executed with the equipment provided in a regular PE gymnasium, sports facility or preschool facility (AST, CAMSA, DEMOST-PRE, FMS-Polygon and ZNA). For all the other tools a manual and specialised test kit are needed (BOT-2, KTK, M-ABC-2 MOBAK, MOT-2 and TGMD-2).

The measurement tools that consist of isolated items are more time consuming than obstacle courses (AST and CAMSA). The obstacle courses take about 15 seconds up to 1 minute to measure one child in contrast to 20 minutes up to 40 minutes per child with a tool that measure tasks separately (BOT-2, DEMOST-pre, KTK, M-ABC-2, MOBAK, MOT-2, TGMD-2 and ZNA).

The costs of the measurement tools range from no costs up to \$ 2045,00. The AST, CAMSA, MOBAK and ZNA are open source. The costs of BOT-2, KTK, M-ABC-2, MOT 4-6, and TGMD-2 range from \$133,00 up to \$ 2045,00. For all the measurement tools a comparable open space, with a flat floor (safe for running) is required to execute the test.





General information			Feasibility					Reference				
Instrument	Developed in	Test goal	Outcome	Test design	Assessor	Product/process	Age (years)	Time	Equipment	Cost	Space needed	Studies (n)
M-ABC-2	1992	To assess motor development and identification of impairment in motor function in children in daily life	Raw item scores are obtained from a set of 30 items. The items are assigned standard scores and percentiles are provided for the three age bands. The total score is the sum of all item scores. The total score is often used for classification purposes based on percentiles	3 age bands (Consist of 8 tasks for each age range in three 5-minute sessions) (1) Fine Motor Aiming & Catching, (2) Fine Motor Skills (Picking Pegs, Threading Laces, and Drawing Trail), (2) Aiming & Catching, which consists of 2 items that consist of 3 items that test gross motor skills (One-shoulder and Throwing, Bouncing onto Mat), and (3) Balance, which consists of 3 items that test gross motor skills (One-shoulder Mat), Walking, Feet-to-toe, Forward, and Hopping on Mats)	Psychologists physiotherapists occupational therapists nurses	Product	3–16	20–40 minutes to assess 1 individual child	Manual of the test and a set of equipment (2) Peg Board, Table, Top Mat, Cubes & Lace Set, Bank Box & Coins, Living Board and Lace Set, Throwing Ball, Bouncing Mat, and Triangle Test (2), Floor Mats in Case (6)	\$ 2045.00	Open space with suitable mats for running and changing (2 x 8 x 10 m)	1, 4, 5, 8, 20, 25, 26, 27, 28, 30, 31, 32, 33, 34, 35
MORAK	2015	To assess basic motor skills and gross motor function in handicapped children	The evaluation can be carried out in several ways. The test assesses "motor movement" as well as using a "motor movement" scale. The total score is easily made by the formation of total values. Through simple raw total values and compared with the total values	4 tasks self-movement (balancing, rolling, jumping, running) and 4 tasks object movement (throwing, catching, bouncing, dribbling)	Sports- and PE teachers	Product	6–10	15–20 minutes to assess 1 individual child	Manual of the test and some basic equipment (1) Swedish walking box, 1 bench, 1 football, 3 basketballs, 1 tennis ball, 1 volleyball, 1 small basketball, 1 rounders, 1 round target	Open source	a regular mat (2 x 8 x 10 m)	11, 26
MOT 4.6	1987	Assesses motor developmental status at 4 to 6 years of age to identify motor disorders	Raw scores, such as the number of jump made, the seconds taken to complete a task, are converted into a three-level ranking scale. The scores of all assessment tasks are then summed to give a total raw motor score ranging between "0" and "34". Converted into a motor score (motor quotient - MQD)	18 items, Forward jump in a hoop (this item is introductory and thus is not assessed), Walking forward, Making dots on a grid, Carrying a dropped stick, Carrying balls from box to box, Catching in backward direction, Throwing a ball to a target, Throwing a ball to a target, Catching a ball on one foot standing on one leg, Catching a ring, Jumping jacks, Jumping over a cord, Rolling around the length of a mat, Rolling a ball up holding a ball on the head, Jumping and landing in a hoop	Preschool professional	Product	4–6	15–20 minutes to assess 1 individual child	Manual of the test and Test kit with specialized materials, which consists of different sports equipment and everyday objects.	\$ 664.00	Open space with flat floor that is suitable for running and changing directions (6 x 4 m)	25, 31
TGM-2	1985	A process-oriented test that was developed to assess the qualitative aspects of motor function in children who are behind peers	Qualitative based Standard scores, based on the Motor Quotient based on percentiles	two subtests: one that assesses performance of locomotion skills (running, sliding, and galloping) and the second subtest assesses object control skills (throwing, catching, kicking, striking, dribbling, and midair hand rolling)	Psychologists physiotherapists PE teachers	Process	3–10	15–20 minutes to assess 1 individual child	Manual of the test and some basic equipment (1) 20–25 cm playground ball, 15–20 cm sponge ball, tennis ball, 20–25 cm plastic or slightly deflated playground ball	\$ 133.00	½ of a regular mat (14 x 10 m)	2, 7, 12, 13, 16, 19, 21, 24, 36, 35
ZMA	2001	To assess the nature range of motor skills and performance in typically developing children from 3;0 to 12 years old	The report of several motor tasks (10) and the motor quotient of movement on a 5-point scale	five components: pure motor skills, fine motor, adaptive, dynamic balance, static balance, and uncoordinated movements	Clinicians, ZMA tester preschool professional	Product and Process	3–12	Total time to perform the test is 20 minutes and 30 minutes, depending on the proficiency of the child	Carton connected with apparatus common to the preschool setting	Open source	Open space with suitable mats for running and changing directions (15 x 5 m)	15, 38

### 7.3.2 Reliability

Of the 38 selected studies, 11 examined the internal consistency, 18 the intra-rater or test-retest reliability and 12 looked into the inter-rater reliability of 8 of the 11 measurement tools (see also Appendix 6).

None of the included studies reported on the reliability of the KTK, MOBAK and the ZNA. Good results have been found on one or more aspects of reliability of the AST, BOT-2, CAMSA, DEMOST-pre, FMS-polygon, and the MOT 4-6 (article 3, 10, 17, 25). With respect to the AST, two aspects of reliability have been studied. The internal consistency of the AST was good and ranged from  $\alpha = 0.700$  up to  $\alpha = 0.764$  (article 14). The test-retest reliability of the AST was also rated as good, ICC ranged from 0.800 (95% CI: 0.669–0.871) up to 0.891 (95% CI: 0.870–0.908) in two studies (article 9, 14). Comparable results were found for the MOT 4-6. The internal consistency of the MOT 4-6 was good ( $\alpha = 0.801$ ) (article 25). The test-retest reliability was also found to be good (ICC = 0.87) (article 25). The test-retest reliability of the DEMOST-pre was also found to be high (ICC: 0.94-0.95) (article 17). The inter-rater reliability was also investigated and was rated as good (ICC = 0.90) in the same study (article 17). Three measurement tools were investigated for only one aspect of reliability (FMS-polygon, CAMSA, BOT-2) (article 29, 3, 10). For both the FMS-polygon and the CAMSA the test-retest reliability was good (FMS-polygon: ICC = 0.98 (article 29); CAMSA: ICC total = 0.91, ICC isolated time score = 0.800, ICC isolated skill score = 0.85 (article 10)). The internal consistency of the BOT-2 was good ( $\alpha = 0.87$ ) for all items (article 3).

Two measurement tools showed varying results considering reliability in the included studies (M-ABC-2, TGMD-2, see Table 7.3). The internal consistency of the M-ABC-2 varied per study. In three studies the internal consistency was moderate (article 1, 28, 23) and in two studies it was rated good (article 20, 27). The ICC values for test-retest reliability ranged from 0.23 up to 0.98 in seven studies (article 34, 37, 22, 27, 28, 20, 23). The inter-rater reliability showed an ICC ranging from ICC = 0.35 up to ICC = 0.99 in five studies (article 37, 22, 27, 20, 23) and the kappa coefficients ranged from 0.95 up to 1.00 in one study (article 33). The internal consistency of the TGMD-2 ranged from  $\alpha = 0.50$  up to  $\alpha = 0.87$  in three studies (article 7, 16, 19). The test-retest reliability of the TGMD-2 was good and ranged from  $r = 0.83$  up to  $r = 0.91$  in four studies (article 7, 13, 16, 24). The intra- and inter-rater reliability coefficients were also rated as good (intra-rater reliability:  $r = 0.92$  up to  $r = 0.99$  (article 24); inter-rater reliability: ICC = 0.71- 0.92, (article 7, 13, 16, 24).

### 7.3.3 Validity

Of the 38 selected studies, 15 examined the construct validity, 2 studies looked into the content validity, and 19 studies examined the concurrent validity (see also Appendix 7). None of the studies investigated elements of validity with regard to the MOT 4-6. For the other ten measurement tools, research was carried out into one or more aspects of validity.

The only measurement tool that is investigated for all three aspects of validity is the M-ABC-2 (article 5, 30, 8, 26, 34, 4, 1, 28, 20, 23). The construct validity of the M-ABC-2 showed a good fit of a 6-item model with values ranging from 0.34 up to 0.77 and Confirmatory Factor Analyses (CFA) confirmed the theoretically assumed 6 factor structure in six studies (article 5, 30, 8, 1, 28, 23). The concurrent validity showed a low to moderate correlation between M-ABC-2 and the TGMD-2 ranging from  $r = 0.13$  up to  $r = 0.40$ . Between the M-ABC-2 and the PDMD-2 a Pearson correlation was found of  $r = 0.631$ . Between M-ABC-2 and BOT-2 Pearson correlations ranged from  $r = 0.60$  up to  $r = 0.90$  in four studies (article 26, 34, 4, 20). Spearman's rho correlation showed no significant correlation between the M-ABC-2 and BOT-2 in one study (article 23). The content validity was investigated in two studies and showed strong Content Validity Index (CVI) concordance results among experts in both studies (article 20, 23).

Two measurement tools have been investigated on their construct and concurrent validity (BOT-2, TGMD-2, article 3, 4, 18). The construct validity of the BOT-2 showed three factors with eigenvalues higher than 1.0, 54.14% score variance in one study (article 3). The concurrent validity can be valued as good and showed a Spearman Rho of 0.80 between the BOT-2 and the M-ABC-2 in one study (article 18) and a low to moderate correlation between BOT-2 short form and the KTK ranging between  $r = 0.30$  and  $r = 0.64$  in one study (article 4). The construct validity of the TGMD-2 showed support of the two-factor model in five studies (article 35, 16, 7, 24, 12). The concurrent validity showed a low to moderate correlation between the TGMD-2 and the KTK ( $r = 0.34 - 0.52$ ), CMAT ( $r = 0.34 - 0.44$ ), and M-ABC-2 ( $r = 0.13-0.40$ ) in four studies (article 26, 2, 19, 24).

For the MOBAK and the DEMOST-pre only the construct validity was investigated (article 17, 11, 36). The construct validity of the DEMOST-pre showed a good model fit (Kaiser-Meyer-Olkin index of 0.85 of two components explaining 54.1% of the variance in one study (article 17)). Also, the two-factorial structure of the MOBAK resulted in a good model fit ( $\chi^2 = 27.56$ ;  $df = 419$ ;  $p = 0.09$ ;  $CFI = 0.97$ ;  $TLI = 0.96$ ;  $RMSEA = 0.037$ ) in two studies (article 16, 19).

Five measurement tools have only been investigated on the concurrent validity in the included studies (AST, CAMSA, FMS-Polygon, KTK, and ZNA, article 9, 14, 10, 29, 2, 6, 18, 15). The concurrent validity of the FMS-polygon was rated as good and showed a correlation between FMS-polygon and the TGMD-2 of  $r = -0.82$  in one study (article 29). The concurrent validity of the AST showed a moderate to good correlation between AST and the KTK ranging from  $r = -0.501$  up to  $r = -0.747$  in two studies (article 9, 14). The concurrent validity of the ZNA was also rated moderate to good based on the correlation between ZNA and the M-ABC-2 ranging from  $r = 0.56$  up to  $r = 0.77$  in one study (article 15). The concurrent validity of the CAMSA showed a moderate correlation with a Spearman Rho ranging from 0.60 up to 0.68 between the CAMSA and the Victorian FMS Assessment in one study (article 10). The concurrent validity of the KTK scored lowest and was rated as poor to moderate (correlation KTK-TGMD-2:  $r = 0.34-0.52$ ; correlation KTK-BOT-2:  $r = 0.25-0.64$ ) in three studies (article 2, 6, 18).

#### 7.3.4 Comparative ratings

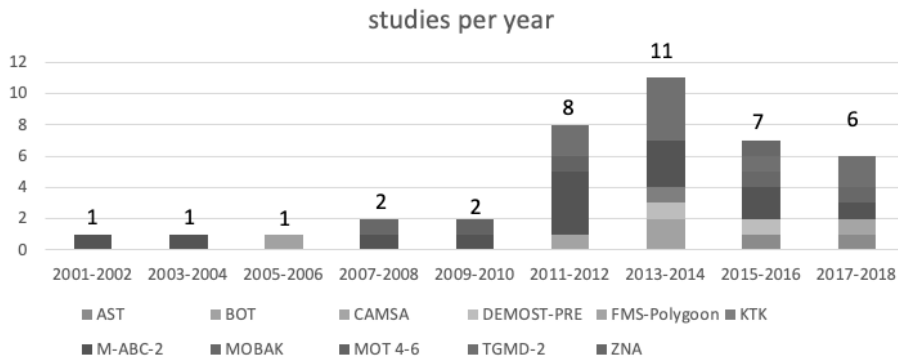
When the outcomes of the 38 studies were weighted for the quality of the study and summarized per measurement tool, it is shown that there is the strongest evidence for the reliability and validity outcomes of the M-ABC and the TGMD-2 (see Table 7.3). The M-ABC scores moderate to good on all items of reliability and validity. There is strong evidence for a good inter-rater reliability (+++), test-retest reliability (+++), and concurrent validity (+++), and there is strong evidence for a moderate internal consistency (±±±) and concurrent validity (±±±) of the M-ABC. With respect to the TGMD-2 there is strong evidence for a good internal consistency (+++), test-retest reliability (+++), and construct validity (+++) and there is strong evidence for a moderate inter-rater reliability (±±±). However, there is also strong evidence for a poor concurrent validity of the TGMD-2 (---). With respect to the AST there is modest evidence for a good internal consistency (++) , test-retest reliability (++) , and concurrent validity (++) of this measurement tool. For all the other measurement tools there is limited evidence for the outcomes (+/±). Especially with respect to the reliability and validity of the DEMOST-pre, KTK, and MOBAK evidence is limited since only few aspects have been investigated in a limited amount of studies and/or in studies of poor quality.

Table 7.3 Evidence for the reliability and validity outcomes

Instrument	Reliability			Validity		Studies Number
	Internal consistency Cronbach's $\alpha$ ( $> 0.60 = +$ , $0.60 - 0.70 = ++$ , $> 0.70 = +++$ )	Intra-Rater Reliability ICC (kappa) ( $< 0.50 = -$ , $0.50 - 0.80 = ++$ , $> 0.80 = +++$ )	Inter-rater Reliability ICC ( $< 0.50 = -$ , $0.50 - 0.80 = ++$ , $> 0.80 = +++$ )	Construct R (Values $< 0.50 = -$ , $0.50 - 0.80 = ++$ , $> 0.80 = +++$ )	Concurrent R ( $< 0.50 = -$ , $0.50 - 0.70 = ++$ , $> 0.70 = +++$ )	
AST	++	++			++	9, 14
BOT	+			±		4, 3, 18
CAMSA	+	+			±	10
DEMOST-PRE		+				17
FMS-POLYGOON		+			+	29
KTK						2, 6, 18
M-ABC-2	±±±±	+++	+++	+++	±±±±	1, 4, 8, 20, 22, 23, 26, 27, 28, 30, 31, 32, 33, 34, 37
MORAK				+		11, 36
MOT 4.6	+	+				23, 31
TGM2	+++	+++	±±±±	+++	---	2, 7, 12, 13, 16, 19, 21, 24, 26, 35
ZNA		±	±		+	13, 38

## 7.4 Discussion

The purpose of the present study was to systematically review the evidence for the reliability, validity, and feasibility of tools to assess children's MC (aged 2-to 18-years old). The search resulted in 38 clinimetric studies on 11 MC measurement tools. In the time slot of this review (2000-2018), the number of studies published on the clinimetric properties of MC measurement tools has grown considerably. In 2001-2002 only one study on this topic has been published in contrast to 11 in 2013-2014. This could possibly be explained by the growing interest in the MC of children (Logan, Ross, Chee, Stodden, & Robinson, 2018) (see Figure 7.2: table of studies published per year).



**Figure 7.2** Overview of studies per year.

In the 38 studies all aspects of reliability have been studied (internal consistency:  $n = 11$ ; test-retest reliability:  $n = 18$ ; and inter-rater reliability:  $n = 12$ ). Also, three aspects of validity have been studied, especially construct ( $n = 15$ ) and concurrent ( $n = 19$ ) have been investigated quite regularly. Although the concurrent validity was investigated in several studies, there is no golden standard in MC measurement (Piek, Hands, & Licari, 2012) and therefore the comparative tests in the included studies differ a lot. The content validity is only examined in two studies (20, 23) by expert opinions, which is considered an adequate method to evaluate the content validity (Polit & Beck, 2006). Although the content validity is often considered to be the most important measurement property of an assessment tool, content validity has not been investigated in many studies because of the challenge of assessing it (Terwee et al., 2018). The content validity concerns the degree to which a sample of items, taken together, constitute a representative definition of a construct (Polit & Beck, 2006). With regard to the content validity, all items in a MC measurement tool should be relevant for the construct of interest

(Patrick et al., 2011). The process-orientated measurement tools are considered more difficult to be evaluated in its' content validity due to many variations in analysing the qualitative outcome (Terwee et al., 2018). Since lack of content validity can affect all other measurement outcomes, researchers are encouraged to evaluate this aspect of validity. These studies might also help to get more insight in de differences between process- and product-orientated MC measurement tools.

In this review five aspects of feasibility have been evaluated in the review process; the assessor, test duration, equipment needed, cost of the measurement tool, and space needed to execute the measurement (see Table 7.2). It is remarkable that only a limited amount of the included studies made statements about the feasibility of the investigated MC measurement tools. The feasibility depends on the purpose of the measurement and the setting in which you want to use the measurement tool. For example; the AST and CAMSA are the only tools that are designed to fit in the PE setting with the purpose to measure motor skill competence of all children.

When looking at the age bands of the participants of the included studies it is shown that pre-schoolers (age 2-4 years) are not often investigated. Out of the 38 studies 10 studies included 3-year-old children. In these studies, less than 15% of the sample was 3-year-old. This might be an important focus for future studies because of the relevance of motor skill development in this age group (Clark & Metcalfe, 2002). On the other hand, the results on the clinimetric properties of the MC measurement tools did not seem to differ between studies with younger versus older children.

When looking at the results of this review it shows that the M-ABC is the only instrument for which there is strong evidence for a moderate to good reliability and validity. The TGMD-2 also scores moderate to good on its' internal consistency, test-retest reliability, construct validity, and inter-rater reliability. However, there is also strong evidence for a poor concurrent validity of the TGMD-2. The M-ABC and TGMD-2 have been investigated in respectively 16 and 10 studies. This might be due to the fact that the TGMD-2 (1985) and the M-ABC (1992) are existing longer than the other measurement tools. Measurement tools that also exist longer than five years i.e., BOT-2 (1978), KTK (1974), MOT 4-6 (1987) and ZNA (2001) have only a 1-character rating that indicates a limited level of evidence for the reliability and validity of the measurement tool. If we look into the measurement tools that have been developed in the last 5 years, such as the AST (2015), CAMSA (2015), DEMOST-pre (2014), FMS-polygon (2011) and MOBAK (2015), it shows that the AST is the only measurement tool that has been rated with two characters. i.e. modest evidence has been found for a good internal consistency, test-retest reliability, and concurrent validity of the AST. The other measurement tools that



have been investigated in the last five years all have a 1-character rating. In the future stronger evidence might be found for these measurement tools.

There are several differences between the older measurement tools and the more recently developed ones. First, the older measurement tools (i.e., TGMD-2, KTK) aim at identifying individuals with mild to severe motor coordination problems, while the more recently developed measurement tools (i.e., CAMSA, AST) aim at assessing motor competence in general, and screen all individuals.

Secondly, in contrast to more recently developed tests that are based on an obstacle course with concatenated motor skills, older measurement tools focus on measuring isolated FMS (Clark & Metcalfe, 2002). The origin of measuring isolated FMS was introduced by (Wickstrom, 1977). Wickstrom (1977) described FMS as basic motor activities that serve as basis for more advanced and sport-specific motor activities. FMS were traditionally categorized in locomotor skills (e.g., jump, hop, and run) and manipulative skills (e.g., throw, kick, and strike) (Fleishman, Quaintance, & Broedling, 1984). In more recent years, researchers defined FMS broader and added balance/stability skills (Barnett et al., 2016; Gallahue, ). Therefore, researchers include balance skills (e.g., balance and body rolling) in the measure of motor competence (Rudd et al., 2015).

Thirdly, from the more recently developed measurement tools three out of five measurement tools are based on an obstacle course which consist of a series of concatenated FMS providing a tool that is likely to more closely mimic how those FMS appear in real life. This idea is based on the concept that motor coordination (e.g., coupling, spatial orientation, and balance ability) is crucial in developing motor competence (Wormhoudt, Teunissen, & Savelsbergh, 2012). Motor coordination can be defined as the ability to have body segments work together in an organized manner (Turvey, 1990) and are considered an underlying component of FMS (Barnett et al., 2016). Tidén et al. (2015) state that one of the shortcomings of traditional motor competence measurement tools is the fact that they have a low ecological validity. Since the tasks in obstacle courses, such as the CAMSA and AST, are completed one after another and change according to the various constraints of the environment, the ecological validity of these obstacle course-based measurement tools might be higher than that of other motor competence tests. However, this has not been investigated yet. Future studies are encouraged to focus on the ecological validity of MC measurement tools (Henderson et al., 2007; Josman, Goffer, & Rosenblum, 2010; Kirby & Sugden, 2007).

This study has some limitations. As the search was restricted to studies published in English, instruments described in other languages were not included in this study. Also, other MC

measurement tools such as questionnaires and measurements tools that focus on perceived motor competence were not included in this study. Since the results of the included studies were all quite positive, the question that can be asked is whether there is publication bias and negative results have not been published.

In conclusion, this review shows there is quite some evidence for a good reliability and validity of several MC measurement tools for 2- to 18-years old children. However, the level of evidence differs per measurement tool with most evidence for the M-ABC and the TGMD-2. It depends on the target group, the purpose and the context which MC measurement tool can best be used.

## 7.5 References

- Altman, D. (1991). *Altman, DG, practical statistics for medical research 1.*
- Barnett, L. M., Stodden, D., Cohen, K. E., Smith, J. J., Lubans, D. R., Lenoir, M., Dudley, D. (2016). Fundamental movement skills: An important focus. *Journal of Teaching in Physical Education, 35*(3), 219-225.
- Barnett, L. M., Van Beurden, E., Morgan, P. J., Brooks, L. O., & Beard, J. R. (2009). Childhood motor skill proficiency as a predictor of adolescent physical activity. *Journal of Adolescent Health, 44*(3), 252-259.
- Branta, C., Haubenstricker, J., & Seefeldt, V. (1984). Age changes in motor skills during childhood and adolescence. *Exercise and Sport Sciences Reviews, 12*, 467-520.
- Bruininks, R. H. (1978). *Bruininks oseretsky test of motor proficiency.* Circle pines-Minnesota: American Guidance Service.
- Bruininks, R. H. (2005). *Bruininks-oseretsky test of motor proficiency* AGS Publishing Circle Pines, MN.
- Clark, J. E., & Metcalfe, J. S. (2002). The mountain of motor development: A metaphor. *Motor Development: Research and Reviews, 2*(163-190).
- Cools, W., Martelaer, K. D., Samaey, C., & Andries, C. (2009). Movement skill assessment of typically developing preschool children: A review of seven movement skill assessment tools. *Journal of Sports Science & Medicine, 8*(2), 154-168.
- Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika, 16*(3), 297-334.
- De Vries, S., Van Hirtum, H., Bakker, I., Hopman-Rock, M., Hirasing, R., & Van Mechelen, W. (2009). Validity and reproducibility of motion sensors in youth: A systematic update. *Medicine Science in Sports Exercise, 41*(4), 818.
- Evers, A., Sijtsma, K., Lucassen, W., & Meijer, R. R. (2010). The dutch review process for evaluating the quality of psychological tests: History, procedure, and results. *International Journal of Testing, 10*(4), 295-317.

Fairclough, S., Stratton, G., & Baldwin, G. (2002). The contribution of secondary school physical education to lifetime physical activity. *European Physical Education Review*, 8(1), 69-84.

Fleishman, E. A., Quaintance, M. K., & Broedling, L. A. (1984). *Taxonomies of human performance: The description of human tasks*. Academic Press.

Gallahue, D. L., & Donnelly, F. C. (2007). *Developmental physical education for all children*. Human Kinetics.

Gallahue, D. L., Ozmun, J. C., & Goodway, J. (2006). *Understanding motor development: Infants, children, adolescents, adults*. Boston.

Gallahue, O. Goodway. (2012). *Development of Fundamental Movement Manipulation Skills*.

Green, K. (2002). Lifelong participation, physical education and the work of ken roberts. *Sport, Education and Society*, 7(2), 167-182.

Henderson, S. E. (1992). *Movement assessment battery for children*. The Psychological Corporation.

Henderson, S. E., Sugden, D. A., & Barnett, A. L. (2007). *Movement assessment battery for children-2*. Harcourt Assessment.

Herrmann, C., Bund, A., Gerlach, E., Kurz, D., Lindemann, U., Rethorst, S., Pühse, U. (2015). A review of the assessment of basic motor qualifications and competencies in school. *International Journal of Physical Education*, 52(3), 2-13.

Hinkle, D., Wiersma, W., & Jurs, S. *Applied statistics for the behavioral sciences*. 2003. Boston, Massachusetts: Houghton Mifflin Boston.

Hoehoer, J., Krijger, M., Savelsbergh, G., & De Vries, S. (2017). Reliability and validity of an athletic skills track to assess motor competence among 4–12-year-old children. *Journal of Science and Medicine in Sport*, 20, e76-e77.

Hoehoer, J., De Vries, S., Krijger-Hombergen, M., Wormhoudt, R., Drent, A., Krabben, K., & Savelsbergh, G. (2016). Validity of an athletic skills track among 6-to 12-year-old children. *Journal of Sports Sciences*, 34(21), 2095-2105.

- Hulteen, R. M., Lander, N. J., Morgan, P. J., Barnett, L. M., Robertson, S. J., & Lubans, D. R. (2015). Validity and reliability of field-based measures for assessing movement skill competency in lifelong physical activities: A systematic review. *Sports Medicine*, 45(10), 1443-1454.
- Josman, N., Goffer, A., & Rosenblum, S. (2010). Development and standardization of a “Do–Eat” activity of daily living performance test for children. *American Journal of Occupational Therapy*, 64(1), 47-58.
- Kambas, A., & Venetsanou, F. (2014). The democritos movement screening tool for preschool children (DEMOST-PRE©): Development and factorial validity. *Research in Developmental Disabilities*, 35(7), 1528-1533.
- Kiphard, E., & Shilling, F. (2007). *Körperkoordinationstest für kinder 2, überarbeitete und ergänzte auflage. beltz test.*
- Kiphard, J., & Schiling, F. (1974). *Korperkoordinations test up kinder. beltz test GmbH.*
- Kirby, A., & Sugden, D. A. (2007). Children with developmental coordination disorders. *Journal of the Royal Society of Medicine*, 100(4), 182-186.
- Kirk, D. (2005). Physical education, youth sport and lifelong participation: The importance of early learning experiences. *European Physical Education Review*, 11(3), 239-255.
- Largo, R. H., Caflisch, J. A., Hug, F., Muggli, K., Molnar, A. A., Molinari, L., Gasser, T. (2001). Neuromotor development from 5 to 18 years. part 1: Timed performance. *Developmental Medicine and Child Neurology*, 43(7), 436-443.
- Logan, S. W., Barnett, L. M., Goodway, J. D., & Stodden, D. F. (2017). Comparison of performance on process-and product-oriented assessments of fundamental motor skills across childhood. *Journal of Sports Sciences*, 35(7), 634-641.
- Logan, S. W., Ross, S. M., Chee, K., Stodden, D. F., & Robinson, L. E. (2018). Fundamental motor skills: A systematic review of terminology. *Journal of Sports Sciences*, 36(7), 781-796.
- Longmuir, P. E., Boyer, C., Lloyd, M., Yang, Y., Boiarskaia, E., Zhu, W., & Tremblay, M. S. (2015). The canadian assessment of physical literacy: Methods for children in grades 4 to 6 (8 to 12 years). *BMC Public Health*, 15(1), 767.
- Lubans, D. R., Morgan, P. J., Cliff, D. P., Barnett, L. M., & Okely, A. D. (2010). Fundamental movement skills in children and adolescents. *Sports Medicine*, 40(12), 1019-1035.

Moher, D., Liberati, A., Tetzlaff, J., & Altman, D. G. (2010). Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *International Journal of Surgery*, 8(5), 336-341.

Okan, O., Pinheiro, P., Zamora, P., & Bauer, U. (2015). Health literacy bei kindern und Jugendlichen Health literacy in childhood and adolescence. *Bundesgesundheitsblatt-Gesundheitsforschung-Gesundheitsschutz*, 58(9), 930-941.

Patrick, D. L., Burke, L. B., Gwaltney, C. J., Leidy, N. K., Martin, M. L., Molsen, E., & Ring, L. (2011). Content validity—establishing and reporting the evidence in newly developed patient-reported outcomes (PRO) instruments for medical product evaluation: ISPOR PRO good research practices task force report: Part 2—assessing respondent understanding. *Value in Health*, 14(8), 978-988.

Piek, J. P., Hands, B., & Licari, M. K. (2012). Assessment of motor functioning in the preschool period. *Neuropsychology Review*, 22(4), 402-413.

Polit, D. F., & Beck, C. T. (2006). The content validity index: Are you sure you know what's being reported? critique and recommendations. *Research in Nursing & Health*, 29(5), 489-497.

Robinson, L. E., Stodden, D. F., Barnett, L. M., Lopes, V. P., Logan, S. W., Rodrigues, L. P., & D'Hondt, E. (2015). Motor competence and its effect on positive developmental trajectories of health. *Sports Medicine*, 45(9), 1273-1284.

Rudd, J. R., Barnett, L. M., Butson, M. L., Farrow, D., Berry, J., & Polman, R. C. (2015). Fundamental movement skills are more than run, throw and catch: The role of stability skills. *PloS One*, 10(10), e0140224.

Schulz, J., Henderson, S. E., Sugden, D. A., & Barnett, A. L. (2011). Structural validity of the movement ABC-2 test: Factor structure comparisons across three age groups. *Research in Developmental Disabilities*, 32(4), 1361-1369.

Terwee, C. B., Prinsen, C. A., Chiarotto, A., Westerman, M., Patrick, D. L., Alonso, J., Mokkink, L. B. (2018). COSMIN methodology for evaluating the content validity of patient-reported outcome measures: A delphi study. *Quality of Life Research*, 27(5), 1159-1170.

- Tidén, A., Lundqvist, C., & Nyberg, M. (2015). Development and initial validation of the NyTid test: A movement assessment tool for compulsory school pupils. *Measurement in Physical Education and Exercise Science*, 19(1), 34-43.
- Turvey, M. T. (1990). Coordination. *American Psychologist*, 45(8), 938.
- Ulrich, D. (2000). TGMD 2–Test of gross motor development examiner’s manual. Austin TX: PRO-ED, 2
- Ulrich, D. A. (1985). TGMD, test of gross motor development Pro-Ed.
- Wickstrom, R. L. (1977). Fundamental motor patterns.
- Wormhoudt, R., Teunissen, J., & Savelsbergh, G. (2012). Athletic skills model. Arko Sports Media, Nieuwegein.
- Zimmer, R., & Volkamer, M. (1987). Motoriktest für vier-bis sechsjährige kinder: Mot 4-6; manual Beltz-Test.
- Zuvela, F., Bozanic, A., & Miletic, D. (2011). POLYGON - A new fundamental movement skills test for 8 year old children: Construction and validation. *Journal of Sports Science & Medicine*, 10(1), 157-163.







# Epilogue

## 8 Epilogue

This thesis describes the development of the Athletic Skills Track (AST). The aim of this thesis is to examine the reliability, validity and feasibility of this new motor competence (MC) assessment tool: The Athletic Skills Track to assess fundamental movement skills among 4- to 12-year-old children in a physical education setting.

There seems to be an urgency to increase our understanding of motor skill development that is also recognised by physical education (PE) teachers. They are willing to monitor motor skill competence of children more objectively (Lander, Morgan, Salmon, & Barnett, 2016). Unfortunately, there seem to be some practical shortcomings in the available assessment tools. The lack of feasibility is, among other things, due to the fact that those are very time-consuming since it takes at least 20 minutes to measure one individual child (Cools et al., 2009). Those shortcomings seem to be the reason many PE teachers currently do not use motor skill competence tests structurally. Therefore, a reliable and valid assessment tool that fits the PE setting is needed.

The main research question can be formulated as follows:

How can physical education teachers assess motor competence to identify motor competence levels of 4-to 12-year old children during a regular physical education lesson in a reliable, valid and feasible manner?

### **Validity of an Athletic Skills Track among 6- to 12-year old children**

In Chapter 2, the feasibility and validity of an Athletic Skills Track (AST) to assess fundamental movement skills among 6- to 12-year-old children in a physical education setting are presented. The AST discussed in this chapter is based on a pilot study that was performed to test if the AST offers opportunities for measuring fundamental movement skills (FMS) with time to complete the AST as solely outcome measurement. In the main study, the AST was refined and validated on a larger scale in a regular PE setting. Four hundred sixty-three Dutch children (211 girls, 252 boys) completed three tests: the Körperkoordinationstest für Kinder (KTK) and two Athletic Skills Tracks (AST-1, AST-2). The validity of AST-1 and AST-2 was examined by correlating the time (s) needed to complete the tracks and the KTK Motor Quotient (MQ).

Overall, there was a low correlation between AST-1 and the KTK MQ ( $r = -0.474$  ( $P < 0.01$ )) and a moderate correlation between AST-2 and the KTK MQ ( $r = -0.502$  ( $P < 0.01$ )). When split up by age group the associations were much higher and ranged between  $r = -0.469$  and  $r = -0.767$ , with the exception of the low correlation coefficient of the AST-2 in 7-year-olds. The results indicate that fundamental movement skills of 6- to 12-year-old children can be assessed with a quick, convenient and low-cost motor competence test in a physical education setting, i.e., an Athletic Skills Track.

### **Reliability and concurrent validity of a motor skill competence test among 4- to 12-year old children**

Chapter 3 describes the investigation of the test-retest reliability, internal consistency and concurrent validity of three Athletic Skills Tracks that were developed based on the AST presented in Chapter 2. Because of the ceiling effect that was found in the first version of the AST (see Chapter 2), three age-related tracks were developed to measure MC of children aged 4- to 12-year old in the PE setting. During a regular PE lesson, 930 4- to 12-year old children (448 girls, 482 boys) completed two motor skill competence tests: (1) the Körperkoordination-Test für Kinder (KTK) and (2) an age-related version of the AST (age 4–6 years: AST-1, age 6–9 years: AST-2, and age 9–12 years: AST-3). The test-retest reliability of the AST was high (AST-1: ICC = 0.881 (95% CI: 0.780–0.934); AST-2: ICC = 0.802 (95% CI: 0.717–0.858); and AST-3: ICC = 0.800 (95% CI: 0.669–0.871)). The internal consistency, concerning the three age-bands of the AST was above the acceptable level of Cronbach's  $\alpha > 0.70$  (AST-1:  $\alpha = 0.764$ ; AST-2:  $\alpha = 0.700$ ; and AST-3:  $\alpha = 0.763$ ). There was a moderate to high correlation between the time to complete the AST, and the age- and gender-related motor quotients of the KTK (AST-1:  $r = -0.747$ ,  $p = 0.01$ ; AST-2:  $r = -0.646$ ,  $p = 0.01$ ; and AST-3:  $r = -0.602$ ,  $p = 0.01$ ). The Athletic Skills Track is a reliable and valid assessment tool to assess motor skill competence among 4- to 12-year old children in the PE setting.

### **The Athletic Skills Track: Age- and gender-related normative values of a motor skills test for 4- to 12-year-old children**

After the conclusion in Chapter 3 that the AST is a reliable and valid assessment tool to assess motor skill competence the next step was to develop normative values belonging to the AST to be able to interpret the individual score of children. In Chapter 4 age- and gender-related normative values belonging to the AST are presented. In a large-scale research design a total

of 7977 Dutch children, 4036 boys and 3941 girls, performed an age-related version of the Athletic Skills Track (AST) as presented in Chapter 3. The AST is a track consisting of 5–7 fundamental movement skill tasks that should be completed as fast as possible. The children performed the test during a regular PE lesson under the supervision of their own PE teacher. For each version of the AST (AST-1:  $n = 917$ ; AST-2:  $n = 3947$ ; AST-3:  $n = 3213$ ) age- and gender-related reference centiles were derived from the gathered data using the Lambda, Mu, Sigma (LMS) method. All children completed the AST within 60 s (mean 29.6 s, SD 7.7). An independent samples t- test showed that boys were significantly faster in completing the track than girls, except for the 4- year-old boys. Therefore, age- and gender-related reference centiles were derived. The reference curves demonstrate an almost linear decrease in time to complete AST-1 and AST-2 with increasing age. The presented study in Chapter 4 provides age- and gender-related normative values and MQ values for the AST among 4- to 12-year-old Dutch children. With these normative values PE teachers can interpret children's performance on the AST.

### **Children's enjoyment of a motor skill test in Physical Education**

After the research about the AST itself in Chapter 2, 3 and 4, the focus in Chapter 5 is a description of a study design to examine children's enjoyment when being tested with the AST in a regular PE lesson. Testing in PE can be valuable, but literature shows that it is important the assessment is a positive experience for all children. The study was conducted among 239 children, aged 4-to-12-years-old. All children completed the AST, where after they rated their enjoyment using a Smileyometer. One week later, 131 children completed another motor skill test, the Körperkoordinationstest für Kinder (KTK), where after they were asked to rank their enjoyment with the AST, the KTK, measurements of body height and weight, and a periodical cognitive test, using a Funsorter. The majority (98%) of the children rated their enjoyment of the AST as good to brilliant. 76% of the children ranked the AST as the most enjoyable in relation to the other three tests. No significant differences were found in enjoyment of the AST between boys and girls, nor between children with different motor competence levels. Most of the children enjoyed performing the AST in a regular PE lesson.

### **Validity and feasibility of an obstacle course to assess fundamental movement skills in a pre-school setting**

Chapter 6 is the result of a collaboration with the University of South Australia in Adelaide. In this chapter the results are presented of a research that aims on measuring MC in the pre-school setting in Australia. For the validation study sixty-five 3-6-year-old children (25 boys and 40 girls) from five pre-schools across Adelaide, Australia participated. Correlations and linear regression analysis (adjusted for age and gender) were used to investigate the association between the time to complete the AST and the raw score of the Test of Gross Motor Development 2 (TGMD-2). For the feasibility study pre-school staff completed a semi-structured interview regarding the feasibility of the AST. The AST took less than a minute per child and the TGMD-2 around 20 minutes for two children. There was a strong negative correlation ( $r = -0.63$ ,  $P < 0.01$ ) between the AST scores and the TGMD-2 scores. All five staff reported strengths of the AST to be its short administration time, setup and appropriateness. These results suggest that the AST could be a feasible and valid method of FMS assessment in Australian pre-schools.

### **Reproducibility, validity and feasibility of motor competence assessment instruments: a systematic review**

In Chapter 7 an overview of the current status in research about MC assessment is presented. The purpose of the review is to systematically review the reliability, validity, and feasibility of MC measurement tools for children aged 2-to 18-years old. Method: A systematic search of the literature was performed in February 2018 in four databases resulting in 4.299 hits. After screening and backwards tracking 38 studies were included. The included studies were evaluated using a 10-item checklist. Results: There is strongest evidence for the reliability and validity of the M-ABC ( $n = 16$  studies) and the TGMD-2 ( $n = 10$ ). The M-ABC scores moderate to good on all items of reliability and validity. With respect to the TGMD-2 there is strong evidence for a good internal consistency, test-retest reliability, construct validity and a moderate inter-rater reliability. However, there is strong evidence for a poor concurrent validity. With respect to the AST there is modest evidence for a good internal consistency, test-retest reliability and concurrent validity. For the other eight reviewed measurement tools there is limited evidence for the clinimetric outcomes. In conclusion, there is quite some evidence for a good reliability and validity of several MC measurement tools for 2- to 18-years old

children. However, the level of evidence differs. It depends on the target group, the purpose and the context which MC measurement tool can best be used.

### 8.1 Main Conclusion

In sum, this thesis aimed to examine the reliability, validity and feasibility of a new MC assessment tool: The Athletic Skills Track to assess fundamental movement skills among 4- to 12-year-old children in a physical education setting.

Chapter 2 showed that fundamental movement skills of 6- to 12-year-old children can be assessed with a quick, convenient and low-cost motor competence test in a physical education setting, i.e., an Athletic Skills Track in a valid matter. Chapter 3 demonstrated that the Athletic Skills Track is a reliable and valid assessment tool to assess motor skill competence among 4- to 12-year old children in the PE setting. Chapter 4 provides age- and gender-related normative values and MQ values for the AST among 4- to 12-year-old Dutch children. With these normative values PE teachers can interpret children's performance on the AST. Chapter 5 illustrated that there are no significant differences in enjoyment of the AST between boys and girls, nor between children with different motor competence levels. Most of the children enjoyed performing the AST in a regular PE lesson. Chapter 6 suggested that the AST could be a feasible and valid method of FMS assessment among Australian pre-school children. Chapter 7 gives an overview of the reliability, validity and feasibility of existing MC assessments presented in scientific literature between 2000 and 2018.

### 8.2 Directions for future studies

The studies in this thesis have several limitations that also might give direction for future studies. This thesis is devoted to the development of a new assessment instrument to measure MC of children. Looking back on the entire process, a number of challenges and dilemmas have become visible that need to be described and may provide room for further research.

#### **The lack of golden standard**

First of all, the lack of a golden standard is one of the issues that has become clearer during the entire research process. Unlike intelligence and language assessments, not one motor competence test has been identified as a gold standard assessment tool (Piek et al., 2012)

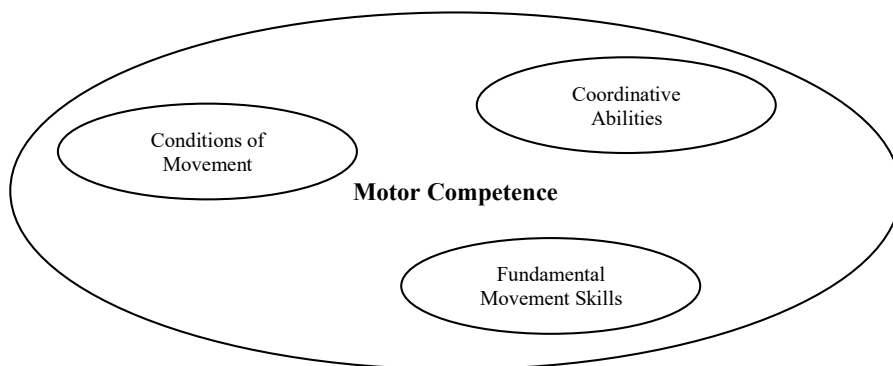
In Chapter 7 eleven MC measurement instruments are described, and it shows that although the concurrent validity was investigated in several studies, the comparative tests in the included

studies differ a lot. Many MC assessment tools produce a total score similar to intelligence tests, which give an indication of the child's overall MC in relation to children of the same age. This implies that there is a single trait for MC which is a controversial issue. Despite the arguments against a total score, Burton and Rodgeron (2001) argue that "the overall composite scores in most MC assessment instruments provide at least rough estimates of MC. Therefore, these tests serve a purpose and future studies should focus on understanding the concept of MC.

### **Measuring Motor Competence?**

Although numerous researches have been carried out in relation to MC measurement tools, the term MC has become confusing because it has been used in various articles as a concept that is defined in different manners (Logan, et al., 2018). Most articles about MC measurement tools take FMS as foundation for measuring MC. These FMS are claimed to provide a foundation for children to develop more specialized movement repertoire, such as sport-specific movement skills for example; making a lay-up in basketball (Clarke & Metcalfe, 2002). FMS are considered as the basis of lifelong movement skills such as swimming and cycling (Hulteen et al., 2015).

When looking at how these FMS are measured in the various MC measuring instruments, it appears that in most measurement tools FMS are measured in an isolated manner. The question is whether such measurement tools measure the total concept of MC. When looking at the presented thesis, it seems possible to stretch the concept of measuring MC by measuring FMS linked to each other. This might give an indication that the total concept of MC is more than the sum of measuring isolated FMS. It seems to raise the question: Will the concept of measuring isolated FMS as a predictor of MC be abandoned in the future and could a more holistic view of MC be introduced?



**Figure 8.1** A conceptual model to display factors on MC.

In Figure 8.1, the model of the development of MC via FMS is further expanded with conditions of movement (i.e. agility, stability, flexibility, endurance and power) and coordinative abilities (i.e. adaptability, balance, coupling, kinetic differentiating, spatial orientation, reaction, rhythmic) based on the ASM model (Wormhoudt et al., 2012). The coordinative abilities were introduced as a part in the development of the AST. In future studies, it might be possible to examine whether this model is closer to defining the measurement of MC than the measurement solely based on isolated FMS. This could perhaps lead to more insight into MC with instruments with a higher ecological validity and maybe even a golden standard.

### **Ecological validity**

When looking at the newer measurement tools two measurement tools are based on an obstacle course which consists of a series of concatenated FMS providing a tool that more closely mimics how those FMS appear in real life. This idea is based on the concept that motor coordination (e.g., coupling, spatial orientation, and balance ability) is crucial in developing motor competence (Wormhoudt, Teunissen, & Savelsbergh, 2012). Tidén et al. (2015) state that one of the shortcomings of traditional motor competence measurement tools is the fact that they show low ecological validity. Since the skills in obstacle courses are completed one after another and change according to the various constraints of the environment, the ecological validity of these obstacle course-based measurement tools might be higher than that of other motor competence tests. The ecological validity of MC measurement tools is stated to be an important feature of testing (Henderson et al., 2007; Kirby & dus Sugden, 2007; Josman et al., 2010). When looking into this phenomenon first, generalizability of the test results should be



investigated. When a child is assessed in a setting that is not ecologically valid, the test-performance may not reflect the child's performance in real life. A child who feels uncomfortable during the test might perform worse than in the real-life situation. In contrast, a child might perform better during the test than during daily life because of extra focus and guidance and less focus areas. Considering the generalizability of measurement tools, everyday performance is best reflected when assessed in a natural setting (Chaytor & Schmitter-Edgecombe 2003). Second, next to the generalizability of the instrument as a whole, the inclusion of the specific tasks in the measurement tool is important as these should relate to the child's everyday performance (Kvavilashvili & Ellis, 2004; Josman et al., 2010). Therefore, it is important to assess tasks that are performed in a real-life setting, and to assess these in a way that reflects everyday performance. Linde et al. (2013) concluded the lack of taking ecological validity into account in MC measurement tools. Further research on MC measurement should also take ecological validity into account.

### **Content validity**

Although the concurrent validity was investigated in several studies as shown in Chapter 7, the content validity is only examined by expert opinions in two studies that were included in the review, which is considered an adequate method to evaluate the content validity (Polit & Beck, 2006). Although the content validity is often considered to be the most important measurement property of an assessment tool, content validity has not been investigated in many studies because of the challenge of assessing it (Terwee et al., 2018). The content validity concerns the degree to which a sample of items, taken together, constitute a representative definition of a construct (Polit & Beck, 2006). With regard to the content validity, all items in a MC measurement tool should be relevant for the construct of interest (Patrick et al., 2011). The process-orientated measurement tools are considered more difficult to be evaluated in its' content validity due to many variations in analysing the qualitative outcome (Terwee et al., 2018). Since lack of content validity can affect all other measurement outcomes, researchers are encouraged to evaluate this aspect of validity. These studies might also help to get more insight in the differences between process- and product-orientated MC measurement tools.

### **Motor competence as part of Physical Literacy**

In this thesis the focus is on measuring MC in the PE setting. The focus on MC seems important because MC is related to all kind of health outcomes (see Figure 1.2). But MC is not the only

component of developing children to healthy active citizens. The model of physical literacy was introduced by Whitehead (2012). In short it can be described as the motivation, confidence, physical competence, knowledge and understanding to value and take responsibility for maintaining healthy active throughout the life course (Whitehead, 2013). MC is considered an important component of physical literacy (Giblin et al., 2014). Fundamental Movement Skill (FMS) play an important role in the development of MC and therefore can be seen as a way of developing the MC component of physical literacy (Almond, 2013). But MC is only one of the components of physical literacy and therefore it seems important to investigate the other components and relations with MC within the model of physical literacy in future studies.

### 8.3 Practical implications

The aim of this thesis was to examine the reliability, validity and feasibility of a new MC assessment tool: The Athletic Skills Track to assess fundamental movement skills among 4- to 12-year-old children in a physical education setting.

This research has started about 5 years ago with a practical question. Out of the network of PE teachers, related to the The Hague University of Applied Sciences, the question arose how they could get more insight into the MC of their pupils. The research has given substance to this practical question in a scientific manner leading to this thesis. A lot has happened in the past 4 years on the practical side of research. In collaboration with ASM B.V. and 2BASICS this research has led to the development of a registration system. This registration system bears the name MQ Scan. In this system PE teachers can enter the scores of the AST. The norm values presented in Chapter 4 are included in this system. The system gives PE teachers the opportunity to screen children on a certain moment in time and monitor them over a longer period of time with multiple measurements. But they can also evaluate new intervention programs they have implemented in their setting or even benchmark the results of their pupils with other pupils. At the moment more than 250 PE teachers have implemented the MQ Scan in their PE setting.

Next to the development of the AST as presented in this thesis a fellow researcher developed the 4-skills scan (Kernebeek et al., 2018) and another colleague research group in the Netherlands developed the HAN Movement Assessment Battery based on the KTK (Platvoet et al., 2018). The 4-skills test is a more traditional measurement tool that measures 4 isolated tasks (standing-still, jumping-force, jumping-coordination and bouncing-ball). It takes up to 10

minutes to measure one individual child with the 4-skills scan. In the research setting the reliability (Kernebeek et al., 2018a) and validity (Kernebeek et al., 2018b) has been investigated. With an ICC of 0.93 (95% confidence interval; 0.92-0.95) for test-retest reliability and 0.97 for inter-rater reliability the 4-skills test seems reliable (Kernebeek et al., 2018a). Correlations between the 4-Skills scan and the M-ABC-2 were moderate ( $r = 0.56$ ) up to high ( $r = 0.64$ ) and therefore the 4-skills scan seems a valid instrument (Kernebeek et al., 2018b). The HAN Movement Assessment Battery is also a more traditional measurement tool that is based on the KTK test. This test consists of 3 items of the KTK i.e., walking backwards, moving sideways, jumping sideways. In addition, one manipulative skill has been added; an eye hand coordination test item.

In the Dutch PE landscape more and more PE teachers start to use the AST, the 4-skills scan or the HAN movement assessment to screen, monitor, evaluate and benchmark MC in their PE setting. It might be interesting to compare the AST, the HAN Movement Assessment Battery and 4-skills scan in a research setting to find out how they relate and what they add to each other.

#### 8.4 The (Applied) Research process

The thesis has been embedded in the research group Healthy Lifestyle in a Supporting Environment of the The Hague University of Applied Sciences and the VU University Amsterdam. The research group Healthy Lifestyle in a Supporting Environment aims at supporting an active lifestyle by doing research that jointly contributes to vital citizens; to a higher quality of life, to sustainably improved social cohesion between and social participation of citizens from young to old. This should involve collaboration between professionals and researchers. To achieve this, a number of starting points have been drawn up. These starting points always serve as a basic principle for the research projects that are carried out within the research group. The development of the AST is one of these projects which also took these basic principles as a starting point.

#### **Demand-driven**

The first yardstick is the demand driven approach of the research. To make it possible to change something in the society is to listen to the society at first. The goal of this thesis started with the practical question of PE teachers. They wanted to get more insight in MC of their pupils.

Because the research question started in collaboration between research and practice, it became possible to do research with lot of participants. The research presented in this thesis is done in collaboration with more than 100 PE teachers who were directly involved in different phases of the research process.

### **Practice-oriented**

The second basic principle is an exploratory research approach. In recent years, applied research has grown increasingly within the Universities of Applied Sciences. In a vision document from the Association of Universities of Applied Sciences, Andriessen et al. (2014) define practice-oriented or applied research as follows: Research is answering questions that lead to relevant knowledge (Andriessen et al., 2014, p.30). This starting point gives direction to the place of research within the University of Applied Sciences. An exploratory research approach fits well with applied sciences. Exploratory research has a number of characteristics. A first characteristic of exploratory research is that it starts from practice. Subsequently one wonders what concepts can support the practical phenomenon (Brohm & Jansen, 2010). Starting from practice creates an inquiring attitude among students, teachers and professional practice because they are made curious by their passion.

A second characteristic is that within exploratory research six phases are distinguished that do not have to be run through linearly. The less linear approach of research makes it possible to be flexible in the opportunities that arise during the research process. Exploratory research often also fits in well with professional practice because practice often does not allow itself to be caught in a linear process.

A third characteristic of exploratory research is that the focus can still change during the research. From education, students are often asked: What are you researching and why? However, the answers to these questions do not have to be clear from the start of the research. In research, students can grow towards the right focus together with professional practice. The goal is to remain curious about what you are not researching. Often at the end of the research the focus becomes really clear (Brohm & Jansen, 2010).

### **Embedding research into the university**

A third basic principle is to embed the research into the University of Applied Science and keep professionals, students and researchers connected throughout a research program, it is important to embed the research into the education program of future professionals. This is also

the case with the research in this thesis in relation to the AST. In order to realize this, the research is always linked to the research group, the students and the practice. This keeps everyone who is part of or comes into contact with the research connected to each other. In addition, it is very important to use innovation management principles based on an innovation cycle. In developing (innovation) and researching the AST, the AST innovation process is placed next to the AST research process. By setting up interdisciplinary teams, it becomes possible to come up with solutions for practical questions across the boundaries of the field.

## 8.5 References

Andriessen, D., & van Praktijkgericht Onderzoek, L. M. (2014). Praktisch relevant én methodisch grondig? Dimensies.

Brohm, R. & Janssen, W. (2010). Kwalitatief onderzoeken. Eburon Uitgeverij BV.

Burton, A. W., & Rodgerson, R. W. (2001). New perspectives on the assessment of movement skills and motor abilities. *Adapted Physical Activity Quarterly*, 18, 347–365.

Chaytor, N., & Schmitter-Edgecombe, M. (2003). The ecological validity of neuropsychological tests: A review of the literature on everyday cognitive skills. *Neuropsychology review*, 13(4), 181-197.

Clark, J. E., & Metcalfe, J. S. (2002). The mountain of motor development: A metaphor. *Motor Development: Research and Reviews*, 2 (163-190).

Josman, N., Goffer, A., & Rosenblum, S. (2010). Development and standardization of a “Do–Eat” activity of daily living performance test for children. *American Journal of Occupational Therapy*, 64(1), 47-58.

Henderson, S. E., Sugden, D. A., & Barnett, A. L. (2007). *Movement assessment battery for children-2* Harcourt Assessment.

Hulteen, R. M., Lander, N. J., Morgan, P. J., Barnett, L. M., Robertson, S. J., & Lubans, D. R. (2015). Validity and reliability of field-based measures for assessing movement skill competency in lifelong physical activities: A systematic review. *Sports Medicine*, 45(10), 1443-1454.

Van Kernebeek, W. G., De Schipper, A. W., Savelsbergh, G. J. P., & Toussaint, H. M. (2018a). Inter-rater and test–retest (between-sessions) reliability of the 4-Skills Scan for dutch elementary school children. *Measurement in Physical Education and Exercise Science*, 22(2), 129–137.

- Van Kernebeek, W. G., De Kroon, M. L. A., Savelsbergh, G. J. P., & Toussaint, H. M. (2018b). The validity of the 4-Skills Scan: A double-validation study. *Scandinavian Journal of Medicine & Science in Sports*, 28(11), 2349–2357.
- Kirby, A., & Sugden, D. A. (2007). Children with developmental coordination disorders. *Journal of the Royal Society of Medicine*, 100(4), 182-186.
- Kvavilashvili, L., & Ellis, J. (2004). Ecological validity and twenty years of real-life/laboratory controversy in memory research: A critical (and historical) review. *History and Philosophy of Psychology*.
- Van der Linde, B. W., Van Netten, J. J., Otten, B., Postema, K., Geuze, R. H., & Schoemaker, M. M. (2013). Development and psychometric properties of the DCDDaily: a new test for clinical assessment of capacity in activities of daily living in children with developmental coordination disorder. *Clinical rehabilitation*, 27(9), 834-844.
- Logan, S. W., Ross, S. M., Chee, K., Stodden, D. F., & Robinson, L. E. (2018). Fundamental motor skills: A systematic review of terminology. *Journal of Sports Sciences*, 36(7), 781-796.
- Patrick, D. L., Burke, L. B., Gwaltney, C. J., Leidy, N. K., Martin, M. L., Molsen, E., & Ring, L. (2011). Content validity—establishing and reporting the evidence in newly developed patient-reported outcomes (PRO) instruments for medical product evaluation: Part 2—assessing respondent understanding. *Value in Health*, 14(8), 978-988.
- Piek, J. P., Hands, B., & Licari, M. K. (2012). Assessment of motor functioning in the preschool period. *Neuropsychology Review*, 22(4), 402-413.
- Polit, D. F., & Beck, C. T. (2006). The content validity index: Are you sure you know what's being reported? critique and recommendations. *Research in Nursing & Health*, 29(5), 489-497.
- Platvoet, S., Faber, I., De Niet, M., Pion, J., Kannekens, R., Elferink-Gemser, M., & Visscher, C. (2018). Development of a tool to assess fundamental movement skills in applied settings. *Frontiers in Education*, Vol. 3, p. 75.

Terwee, C. B., Prinsen, C. A., Chiarotto, A., Westerman, M., Patrick, D. L., Alonso, J., Mokkink, L. B. (2018). COSMIN methodology for evaluating the content validity of patient-reported outcome measures: A delphi study. *Quality of Life Research*, 27(5), 1159-1170.

Tidén, A., Lundqvist, C., & Nyberg, M. (2015). Development and initial validation of the NyTid test: A movement assessment tool for compulsory school pupils. *Measurement in Physical Education and Exercise Science*, 19(1), 34-43.

Whitehead, M. (2013). Definition of physical literacy and clarification of related issues. *ICSSPE Bulletin*, 65 (1.2).

Wormhoudt, R., Teunissen, J., & Savelsbergh, G. (2012). Athletic skills model. Arko Sports Media, Nieuwegein.









Dutch summary (*Samenvatting*)

Dankwoord

About the Author

List of publications

Scientific conference  
proceedings

Other Media

Awards

Appendix 1, 2, 3, 4, 5, 6, 7 en 8

---

# A

## Samenvatting

Dit proefschrift beschrijft de ontwikkeling en evaluatie van de Athletic Skills Track, een nieuwe motoriektest voor in de lichamelijke opvoeding. In dit proefschrift wordt ingegaan op de betrouwbaarheid, validiteit en praktische toepasbaarheid van deze nieuwe motoriektest.

Kinderen zijn motorisch minder vaardig dan vroeger. Gymleraren erkennen dit probleem en zoeken naar mogelijkheden om de motorische vaardigheid van kinderen te verbeteren. Hiervoor is het nodig de motorische vaardigheid van kinderen objectiever te volgen in de dagelijkse praktijk van de lichamelijke opvoeding. De beschikbare meetinstrumenten zijn echter niet praktisch toepasbaar. Ze zijn tijdrovend, aangezien het minstens 20 minuten duurt om één individueel kind te meten en niet gericht op het meten in de context van de gymles. Deze tekortkomingen lijken de reden te zijn dat veel gymleraren momenteel geen gebruik maken van motoriektesten. Daarom is een betrouwbare, valide en praktisch toepasbare motoriektest nodig die past binnen de context van de gymles.

De belangrijkste onderzoeksvraag kan als volgt worden geformuleerd:

Hoe kunnen gymleerkrachten in het basisonderwijs de motorische vaardigheid beoordelen bij 4 tot 12-jarige kinderen tijdens een reguliere gymles op een betrouwbare, valide en haalbare manier?

### **Validiteit van een Athletic Skills Track bij kinderen van 6 tot 12 jaar**

In Hoofdstuk 2 wordt de praktische toepasbaarheid en validiteit van een Athletic Skills Track (AST) gepresenteerd voor het beoordelen van de motorische vaardigheid van 6- tot 12-jarige kinderen in de gymles. Ruim 450 Nederlandse kinderen (211 meisjes, 252 jongens) voltooiden drie testen: de Körperkoordinationstest für Kinder (KTK) en twee Athletic Skills Tracks (AST-1, AST-2). De validiteit van AST-1 en AST-2 werd onderzocht door de tijd die nodig was om de track te voltooien te correleren met de KTK Motor Quotient (MQ).

Over het algemeen was er een lage correlatie tussen AST-1 en de KTK MQ ( $r = -0,474$  ( $P < 0,01$ )) en een matige correlatie tussen AST-2 en de KTK MQ ( $r = -0,502$  ( $P < 0,01$ )). Bij opsplitsing naar leeftijdsgroepen waren de associaties veel hoger en varieerden deze tussen  $r = -0,469$  en  $r = -0,776$ , met uitzondering van de lage correlatiecoëfficiënt van de AST-2 bij 7-jarigen. De resultaten geven aan dat de motorische vaardigheid van 6 - tot 12-jarige kinderen kan worden beoordeeld met de Athletic Skills Track.

### **Betrouwbaarheid en concurrente validiteit van de Athletic Skills Track: een motorische vaardigheidstest bij 4- tot 12-jarige kinderen**

Hoofdstuk 3 beschrijft het onderzoek naar de test-hertest betrouwbaarheid, interne consistentie en concurrente validiteit van drie Athletic Skills Tracks die werden ontwikkeld op basis van de AST gepresenteerd in Hoofdstuk 2. Vanwege het plafondeffect dat werd gevonden in de eerste versie van de AST (zie Hoofdstuk 2), werden drie leeftijdsspecifieke hindernisbanen, met 5 tot 7 gekoppelde fundamentele motorische vaardigheden, ontwikkeld om de motorische vaardigheid van kinderen in de leeftijd van 4- tot 12-jarigen te meten in een reguliere gymles. Tijdens een gewone gymles voltooiden 930 4- tot 12-jarige kinderen (448 meisjes, 482 jongens) twee motorische vaardigheidstesten: (1) de KTK en (2) een leeftijd gerelateerde versie van de AST (leeftijd 4-6 jaar: AST-1, leeftijd 6-9 jaar: AST-2 en leeftijd 9-12 jaar: AST-3). De test-hertest betrouwbaarheid van de AST was hoog (AST-1: ICC = 0,881 (95% CI: 0,780-0,934), AST-2: ICC = 0,802 (95% CI: 0,717-0,858) en AST-3: ICC = 0,800 (95% CI: 0,669-0,871). De interne consistentie van de drie leeftijd gerelateerde versies van de AST lag boven het acceptabele niveau van Cronbach's  $\alpha > 0.70$  (AST-1:  $\alpha = 0,764$ ; AST-2:  $\alpha = 0,700$  en AST-3:  $\alpha = 0,763$ ). Er was een matige tot hoge correlatie tussen de tijd om de AST te voltooien en de leeftijd- en geslacht gerelateerde motorquotiënten van de KTK (AST-1:  $r = -0,747$ ,  $p = 0,01$ ; AST-2:  $r = -0,646$ ,  $p = 0,01$ ; en AST-3:  $r = -0,602$ ,  $p = 0,01$ ). Uit deze studie blijkt dat de Athletic Skills Track een betrouwbaar en valide meetinstrument is om de motorische vaardigheid van 4- tot 12-jarige kinderen in de gymles te meten.

### **De Athletic Skills Track: leeftijds- en geslacht gerelateerde normatieve waarden van een motorische vaardigheidstest voor 4- tot 12-jarige kinderen**

Na de conclusie in hoofdstuk 3 dat de AST een betrouwbaar en valide meetinstrument is om de motorische vaardigheid te meten, is de volgende stap het ontwikkelen van normatieve waarden voor de AST, om de individuele score van kinderen te kunnen interpreteren. In Hoofdstuk 4 worden leeftijds- en geslacht gerelateerde normatieve waarden voor de AST gepresenteerd. In een grootschalig onderzoek hebben in totaal 7977 Nederlandse kinderen, 4036 jongens en 3941 meisjes, een leeftijd gerelateerde versie van de AST uitgevoerd. De kinderen voerden de test uit tijdens een gewone gymles onder toezicht van hun eigen gymleraar. Voor elke versie van de AST (AST-1:  $n = 917$ ; AST-2:  $n = 3947$ ; AST-3:  $n = 3213$ ) werden leeftijds- en geslacht gerelateerde referentiewaarden afgeleid van de verzamelde gegevens met behulp van de Lambda, Mu, Sigma (LMS) methode. Alle kinderen voltooiden

de AST binnen 60 seconden (gemiddeld 29,6 seconden,  $\pm 7,7$ ). Een independent samples t- test toonde aan dat de jongens beduidend sneller waren in het voltooien van de AST dan meisjes, behalve de 4-jarige jongens. Daarom werden leeftijds- en geslacht gerelateerde referentiecentielen ontwikkeld. De referentiecurves tonen een bijna lineaire afname in de tijd om AST-1 en AST-2 met toenemende leeftijd te voltooien. Naast de leeftijds- en geslacht gerelateerde normatieve waarden voor de AST zijn ook MQ-waarden berekend vergelijkbaar met IQ-waarden waarmee de tijden op de AST nog sneller te duiden zijn.

### **De beleving van kinderen in relatie tot een motoriektest in de lichamelijke opvoeding.**

Na een drietal onderzoeken naar de AST in Hoofdstuk 2, 3 en 4, ligt de focus in Hoofdstuk 5 op een beschrijving van een onderzoek naar het plezier van kinderen tijdens het voltooien van de AST in een gewone gymles. Het kan waardevol zijn om metingen uit te voeren in de lichamelijke opvoeding, maar eerder onderzoek liet zien dat het belangrijk is dat een test een positieve, plezierige ervaring is voor alle kinderen. Het onderzoek dat in Hoofdstuk 5 wordt gepresenteerd, is uitgevoerd onder 239 kinderen van 4 tot 12 jaar oud. Alle kinderen voltooiden de AST, waarna hun plezier beoordeeld werd met behulp van een Smileyometer. Een week later voltooiden 131 kinderen een andere motoriektest, de KTK. Naderhand werd gevraagd met een Funsorter aan te geven hoeveel plezier zij ervaren tijdens de AST, de KTK, metingen van lichaamslengte en gewicht, en de CITO-toets. De meerderheid (98%) van de kinderen beoordeelden de AST als zeer plezierig. 76% van de kinderen beoordeelde de AST als de meest plezierige van de vier testen. Er werden geen significante verschillen gevonden in het plezier van de AST tussen jongens en meisjes, noch tussen kinderen met verschillende motorische competentieniveaus. Uit dit onderzoek kan geconcludeerd worden dat de meeste kinderen het uitvoeren van de AST in een normale gymles ervaren als plezierig.

### **Validiteit en praktische toepasbaarheid van de Athletic Skills Track om de motorische vaardigheid in een voorschoolse setting te beoordelen**

Hoofdstuk 6 is het resultaat van een samenwerking met de University of South Australia in Adelaide. In dit hoofdstuk worden de resultaten gepresenteerd van een onderzoek dat gericht is op het meten van motorische vaardigheid in de voorschoolse opvang in Australië. Aan de validatie-studie namen vijftien 3-6-jarige kinderen (25 jongens en 40 meisjes) uit vijf kinderdagverblijven uit Adelaide deel. Correlaties en lineaire regressieanalyse (gecorrigeerd voor leeftijd en geslacht) werden gebruikt om de associatie tussen de tijd om de AST te

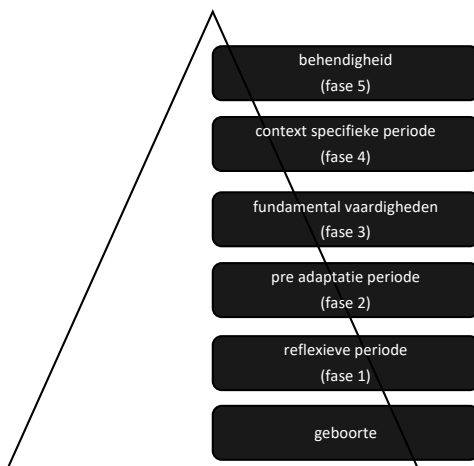
volbrengen en de onbewerkte score van de Test of Gross Motor Development 2 (TGMD-2) te onderzoeken. Daarnaast is onder het personeel van de kleuterschool een semigestructureerd interview afgenomen over de haalbaarheid van de AST. De AST duurde minder dan een minuut per kind en de TGMD-2 ongeveer 20 minuten voor twee kinderen. Er was een sterke negatieve correlatie ( $r = -0,63$ ,  $p < 0,01$ ) tussen de AST-scores en de TGMD-2-scores. Alle vijf medewerkers meldden als sterke punten van de AST; de korte afdruk tijd, eenvoudige opzet en geschiktheid voor de doelgroep. Deze resultaten suggereren dat de AST een praktisch toepasbare en valide methode zou kunnen zijn voor de beoordeling van motorische vaardigheid van 3-6-jarige kinderen op Australische kinderdagverblijven.

### **Betrouwbaarheid, validiteit en praktische toepasbaarheid van motorische vaardigheidstesten: een systematische review**

In Hoofdstuk 7 worden de resultaten gepresenteerd van een systematische review naar de klinimetrische kwaliteit van de verschillende motorische vaardigheidstesten zoals in de wetenschappelijke literatuur tussen 2000 en 2018 zijn gepubliceerd. In de systematische review zijn na selectie 38 artikelen beoordeeld die in totaal 11 motorische vaardigheidstesten op verschillende onderdelen van betrouwbaarheid, validiteit en praktische toepasbaarheid onderzocht hebben. Uit de systematische review blijkt dat er behoorlijk wat bewijs is voor goede betrouwbaarheid en validiteit van verschillende motorische vaardigheidstesten. Het bewijsniveau verschilt echter per motorische test. Daarnaast hangt het af van de doelgroep, het doel van de meting en de context waarin gemeten wordt welke motorische vaardigheidstest het best kan worden gebruikt.

## Dankwoord

Wat ontzettend bijzonder dat na 3,5 jaar mijn proefschrift af is. Er is veel gebeurd in die tijd en WE hebben veel werk verzet. Inderdaad “WE”, want als er één ding is dat ik heb geleerd: promoveren doe je niet alleen. Dankzij de samenwerking met en steun van verschillende mensen met allemaal unieke bijdragen aan het onderzoek is dit proefschrift tot stand gekomen. Mijn promotietraject heeft veel fases gekend die, als ik er zo eens over nadenk, gelijkenissen vertonen met de berg van motorisch leren die het uitgangspunt is geweest van dit promotieonderzoek. Deze metafoor in een notendop. **Fase 0.** Geboorte, ineens was ik aan het begin van een promotieonderzoek, NWO keurde mijn promotie aanvraag goed, de officiële



start van mijn promotietraject. Een compleet nieuwe wereld. **Fase 1.** Reflexieve periode; eigenlijk was ik in het begin vanuit reflexen aan het handelen. Ik kon niet meer dan een beetje reflexmatig bijna spastisch reageren, heel puur en onbewust. Toevallig beland in een promotietraject en geen idee hoe te handelen, de wetenschap was compleet nieuw. **Fase 2.** Pre-adaptatie; in het begin was ik vooral praktisch bezig, zoals met het uitvoeren van metingen. Daarnaast heb ik behoorlijk wat inhaalwerk op vaardigheden gedaan in mijn scholingsprogramma rondom statistiek en wetenschappelijk schrijven. **Fase 3.** Fundamentele vaardigheden; langzamerhand kreeg ik het idee dat ik de fundamentele vaardigheden begon door te krijgen. Gaandeweg kwam er meer ruimte om aan het wetenschappelijk gedeelte te werken en kon ik me meer gaan focussen op het schrijven van de artikelen. **Fase 4.** Context specifieke periode; nu de basis redelijk in bezit was, werd het wel zaak om deze toe te passen in verschillende contexten. Op congressen, in overleggen, bij het publiceren en het opzetten van vervolgonderzoeken bleek het steeds eenvoudiger om te schakelen. **Fase 5.** Behendigheid; aan het einde van het traject denk ik dat ik heb uitgevonden wat promoveren inhoudt en besef ik dat:

“The only true wisdom is in knowing you know nothing”  
 (Socrates 469 v.C. – 399 v.C.)”



Vooral omdat wetenschap niet met de dag komt, het lang kan duren voordat je een succes boek heb je mensen nodig. Mensen die je motiveren en het gevoel geven dat je op de goede weg zit en mensen die je ondersteunen met hun vaardigheden en expertise. Gelukkig had ik genoeg mensen om me heen, die me in alle fases hebben geholpen en gesteund. Ik wil iedereen bedanken die bijgedragen heeft aan mijn proefschrift!

Ik heb veel geluk gehad met mijn promotieteam. Al waren we uiteindelijk maar met zijn drieën, dit was voor mij precies genoeg. Aan de ene kant **Geert Savelsbergh** (promotor), de strateeg die mij met beide benen op de grond heeft gehouden. Een uitspraak zal ik nooit vergeten: “Joris, het is geen ‘rocket science’ wat we hier doen, doe maar een beetje rustig.” Hopelijk kunnen we de samenwerking voortzetten, dat zou mooi zijn!

Aan de andere kant **Sanne de Vries** (copromotor), de rots in de branding, die in dit proces een grote bijdrage heeft geleverd om mij te begeleiden tot waar ik nu sta. Naast het afronden van mijn promotie heeft het traject mij nog wat belangrijks opgeleverd: een goede vriendin, waar ik op kan rekenen. Ik hoop dat dit in de toekomst zo kan blijven en dat we nog veel onderzoeksprojecten samen mogen uitvoeren.

Binnen de onderzoeksgroep Gezonde Leefstijl in een Stimulerende Omgeving heb ik ook altijd veel steun ervaren van verschillende mensen. **Michiel Krijger** (paranimf), die altijd in is voor een geintje, mij de eerste stapjes in SPSS en Refworks heeft bijgebracht. Vooral in het begin van dit promotietraject was je van onschatbare waarde. Alleen al omdat we samen cursussen volgden bij EpidM en ons daar kostelijk vermaakt hebben. Ook fijn dat we ook over andere belangrijke dingen zoals golfsurfen kunnen praten. Laten we dat vooral maar voortzetten. **Annemarie de Witte** (paranimf), net 1,5 jaar eerder in het bootje van promoveren gestapt. Dat was voor mij perfect, kon ik lekker bij je afkijken. Tijdens de nodige hardloop trainingen heb je me laten zien en verteld wat wel handig en niet handig is in een promotietraject. Daarnaast heb je ook de lay-out van dit boekje ontworpen. **Michel Bosman** (collega-onderzoeker), vriend en collega die eigenlijk altijd wel in staat is om de spiegel voor te houden of de andere kant te laten zien. Nu dit traject erop zit is het misschien weer eens tijd om samen een mooi project op te pakken, ons boek herschrijven is zeker een goede optie denk ik. **Tinus Jongert** (directeur NPI), dankzij jou ben ik in het lectoraat GLSO terecht gekomen. In de korte tijd die ik binnen het lectoraat heb gewerkt heb ik altijd veel steun en vertrouwen gevoeld. Fijn dat we elkaar nu nog steeds spreken, laten we dat vooral blijven doen.

Daarnaast zijn er binnen het lectoraat Revalidatie drie mensen waar ik gaandeweg een speciale band mee heb opgebouwd. **Monique Berger** (onderzoeker), altijd fijn om met jou te praten en samen nieuwe plannen te maken. Jij noemt mij altijd één van je favoriete collega's en dat betekent veel voor mij en dat is zeker wederzijds. Daarnaast **Arend de Kloet** (lector) en **Jorít Meesters** (onderzoeker) dank voor de mooie tijd in Brazilië: mooie herinneringen die ik niet snel vergeet.

Binnen De Haagse Hogeschool, die als organisatie veel in mij geïnvesteerd heeft, zijn nog veel meer mensen op te noemen die allemaal op de één of andere manier hebben bijgedragen aan dit promotietraject; **Hans Tromp, John van de Berg, Guido Ongena, Sophie Mentink, Ellen Poot, Esther Lankhuijzen, Kees Mol, Jacco van Uden, Ineke van der Meule, Pim Breebaart, Elly Taal, Marianne Nannings, Bert van der Tuuk, Guy van der Meulen en Son Burgers-Geerinckx.**

Maar ook binnen het werkveld zijn er veel mensen die ik wil bedanken. Allereerst **René Wormhoudt**, die als grondlegger van het Athletic Skills Model aan de basis van deze promotie heeft gestaan. We hebben zeker in het begin veel gepraat over motorische ontwikkeling van kinderen en hoe we dit kunnen meten. Vanuit de eerste ideeën is uiteindelijk deze thesis ontstaan. Fijn dat iemand vanuit zijn expertise belangrijke zaken toevoegt. Daar kan ik nog steeds van genieten.

Praktijkgericht onderzoek kan plaatsvinden door onder andere de inzet van het werkveld. Maar door de kwaliteit van het werkveld wordt het onderzoek vervolgens een succes. Lof voor alle gymdocenten in Den Haag en omstreken, docenten van **Stichting de Haagse Scholen** en in het bijzonder **Rob Smit en Jan Kalter** die beiden ervoor zorgde dat het onderzoek altijd langs de lat van de praktijk werd gelegd.

De andere component van praktijkgericht onderzoek zijn de studenten. Samen met een grote groep doortastende, onderzoekende en enthousiaste studenten van de HALO heeft dit onderzoek zo grootschalig kunnen plaatsvinden. De eerste lichte studenten die in dit onderzoek hebben geparticipeerd zijn: **Anneleid, Wendy, Michiel, Jill, Jeroen, Yordi, Jorien, Ilse, Rachelle, Twan, Michelle, Dylan, Patrick, Bob, Marc, Meta, Trudy, Joëlle Erik, Bob, Matthijs, Naomi en Fatih.** Deze groep studenten heeft veel betekend in de beginfase van dit onderzoek. Na de succesvolle pilot binnen het curriculum van de HALO waar deze groep aan deelgenomen heeft zijn er, na deze groep, nog honderden studenten betrokken geweest bij dit onderzoek. In het bijzonder wil ik **Roxanne Marges, Erik Stolk en Jeroen de Groot** benoemen die in het onderzoek een speciale rol hebben gehad.

Promoveren is werk maar in de privésfeer heeft een promotietraject misschien wel de grootste impact. Zowel in voor- als tegenspoed, hebben jullie mij altijd gesteund. **Martine**, mijn grote liefde. Jij hebt mij de ruimte gegeven om dit traject te doorlopen. Omdat de basis stabiel en veilig is kan ik de hele wereld aan. Dat blijkt wel anders was het niet gelukt. **Renske, Nanna** en **Figo**, mijn lieve apen. Jullie zorgen voor een glimlach op mijn gezicht. Jullie bijdrage aan het onderzoek is niet onverdienlijk. Jullie hebben enthousiast geholpen bij het acteren in promotiefilmpjes en als ik weer eens weg was voor onderzoek werd dat eigenlijk altijd geaccepteerd en waren jullie extra lief voor mamma. Leukere kinderen kan een vader zich niet wensen.

**Pappa, mamma** en mijn grote **broer (Edgar)**, die mij van jongs af aan hebben geleerd om ontwikkelings-gericht te denken. Blijkbaar toch een goede keuze geweest om iets met bewegen te gaan doen. Jullie zijn zeer betrokken geweest bij mijn promotietraject. Altijd geïnteresseerd, trots en steunend als dat nodig was.

Mijn **schoonfamilie**, jullie hebben mij altijd verder geholpen in mijn ontwikkeling. **Gijs** die nog even vanuit de Verenigde Staten een artikel van feedback voorziet, **Frederik** die even meedenkt over de te nemen stappen in het proces en mij “the unwritten rules of PhD” heeft bijgebracht. **Sam** en **Margriet** die in Frankrijk een fijne uitvalsbasis hebben geboden, Les Grandes Cheintres à Chaumard, waar ik tijd en ruimte had om artikelen te schrijven en te praten over mijn promotietraject.

Natuurlijk is ook het fijne netwerk van vrienden belangrijk omdat jullie er eigenlijk altijd voor mij zijn. Fijn dat ik met jullie allemaal bevriend mag zijn. Ik zal de komende tijd weer wat meer tijd in jullie investeren (promoveren is verwaarlozing).

Als laatste wil ik mijn dank betuigen aan een veel te vroeg overleden goede vriend **Prof. dr. Roberto Perez**. Als buurman kon ik altijd even bij je langs om te sparren over het onderzoek. Altijd goede adviezen en veel lol. Ik mis je...

Ik ben trots op het resultaat van mijn proefschrift, maar ik realiseer me terdege dat dit niet was gelukt zonder jullie. Bedankt!!

## About the author



Joris Hoeboer was born on the 19<sup>th</sup> of June 1977 in Rozenburg, the Netherlands. In 1999 he graduated from the Bachelor of Physical Education at The Hague University of Applied Sciences. After a period of working as a PE teacher at a secondary school combined with coaching in elite women gymnastics he started as teacher in gymnastics and motor learning at The Hague University of Applied Sciences (THUAS) in 2004. In this period, he wrote a book about gymnastics together with Michel Bosman. In 2012 he finished his professional master Innovation Management at Pro Education in Amsterdam that lay the foundation of the

PhD project of the presented thesis.

In 2014 Joris started as researcher in the research group: Healthy Lifestyle in a Supporting Environment at THUAS supervised by dr. Sanne de Vries. In this research group Joris is one of the researchers in the Gym of the Future research program, where he is responsible for applied research in the field of Motor Learning and Performance. After successfully submitting a grant application, that was approved by National Science Organization (NWO) (9023.006.005) in 2015, he started the PhD-research in 2016. This thesis is the result of his PhD-research.

Email: [j.j.a.a.hoeboer@hhs.nl](mailto:j.j.a.a.hoeboer@hhs.nl)

Phone: 06-89990117

## Publications

Publications as part of this thesis

**Hoeboer, J. J. A. A. M.**, Opstoel, K., Martelaer, K., Savelsbergh, G., & de Vries, S. I. (submitted, December 2018). Reproducibility, validity and feasibility of motor competence assessment instruments: a systematic review.

Brooke M. Klingberg, **Joris J.J.A.A.M. Hoeboer**, Natasha Schranz, Lisa M. Barnett, Sanne I. de Vries, Katia Ferrar (accepted December 2018). Validity and feasibility of an obstacle course to assess fundamental movement skills in a pre-school setting. *Journal of Sports Sciences*, 1-9.

**Hoeboer, J. J. A. A. M.**, Ongena, G., Krijger-Hombergen, M., Stolk, E., Savelsbergh, G. J. P., & de Vries, S. I. (2018). The Athletic Skills Track: Age-and gender-related normative values of a motor skills test for 4-to 12-year-old children. *Journal of Science and Medicine in Sport*, Volume 20, e76 - e77.

**Hoeboer, J.**, Krijger-Hombergen, M., Savelsbergh, G., & De Vries, S. (2018). Reliability and concurrent validity of a motor skill competence test among 4-to 12-year old children. *Journal of Sports Sciences*, 36(14), 1607-1613.

**Hoeboer, J.**, de Vries, S., Mast, D., & Savelsbergh, G. (2017). Children's enjoyment of a motor skill test in physical education. *Journal of Physical Education Research*, Volume 4, Issue IV, 01-16.

**Hoeboer, J.**, De Vries, S., Krijger-Hombergen, M., Wormhoudt, R., Drent, A., Krabben, K., & Savelsbergh, G. (2016). Validity of an Athletic Skills Track among 6-to 12-year-old children. *Journal of Sports Sciences*, 34(21), 2095-2105.

## Other publications

Adank, A.M., Van Kann, D. H. H., **Hoeboer, J. J.**, de Vries, S. I., Kremers, S. P. J., Vos, S. B. (2018). Investigating Motor Competence in Association with Sedentary Behavior and Physical Activity in 7- to 11-Year-Old Children. *International Journal of Environmental Research and Public Health* 15, 11: 2470.

Martelaer, K. de, Opstoel, K., Vries, S. de, **Hoeboer, J.**, Cools, W., Noordstart, J., Lucassen, J., D'Hondt, E., Lenoir, M. (2017). Abstract Challenges for big data use of Fundamental Movement Skill performance in a school setting. In: A Bund, C Scheuer (eds.). *Changes in childhood and adolescence: current challenges for physical education. Proceedings of the 12th FIEP European Congress, 13-16 september 2017, Esch-sur-Alzette, Luxembourg.* Berlin: Logos Verlag Berlin, pp 106-107.

**Hoeboer, J.**, Savelsbergh, G., Vries, S. de (2017). Abstract What do children think of the Athletic Skills Track as a motor competence test in physical education? In: A Bund, C Scheuer (eds.). *Changes in childhood and adolescence: current challenges for physical education. Proceedings of the 12th FIEP European Congress, 13-16 september 2017, Esch-sur-Alzette, Luxembourg.* Berlin: Logos Verlag Berlin, pp 110-111.

**Hoeboer, J.**, Krijger, M., Ongena, G., Stolk, E., Savelsbergh, G., Vries, S.I. de (2017) The Athletic Skills Track: age- and gender-specific normative values of a motor skills test for 4- to 12-year-old children. In: A Bund, C Scheuer (eds.). *Changes in childhood and adolescence: current challenges for physical education. Proceedings of the 12th FIEP European Congress, 13-16 september 2017, Esch-sur-Alzette, Luxembourg.* Berlin: Logos Verlag Berlin, pp 112-113.

Krijger, M., **Hoeboer, J.** Vries, S. de (2017). Abstract Accuracy of PE teachers to estimate children's motor competence. In: A Bund, C Scheuer (eds.). *Changes in childhood and adolescence: current challenges for physical education. Proceedings of the 12th FIEP European Congress, 13-16 september 2017, Esch-sur-Alzette, Luxembourg.* Berlin: Logos Verlag Berlin, pp 140-141.

**Hoeboer J.**, de Vries S.I. (2015). Gymzaal van de Toekomst: samen leren in een innovatieomgeving. ART, 8.

**Hoeboer J.**, Krijger M., Savelsbergh G., De Vries S. (2014). De bijdrage van LO aan de motorische ontwikkeling van kinderen. Lichamelijke opvoeding, 9: 49-51.

**Hoeboer J.**, Krijger M., De Vries S. de. (2014). Motorische ontwikkeling in relatie tot een actieve leefstijl. Sport & Strategie online, <http://www.sportenstrategie.nl/partners/31137/>.

**Hoeboer J.**, Krijger M., De Vries S. de. (2014). Motorische ontwikkeling in relatie tot een actieve leefstijl. Sportgericht, 1 (68): 28-31.

Altenburg T., **Hoeboer J.**, Bosman M., Krijger M., Bakker J., Toussaint H., Savelsbergh G. (2010). Centrale visuele sturing verbetert de bewegingsuitvoering van de radslag. Sportgericht, 2010; 2 (64): 28-31. 46-49.

Bosman M., & **Hoeboer J.** (2009). Turnen, Die Bausteine. GrafiSelect B.V., Haaksbergen.

Bosman M., & **Hoeboer J.** (2008). De Bouwstenen van het turnen op school (2e druk). GrafiSelect B.V., Haaksbergen.

Bosman M., **Hoeboer J.**, Bakker J., Krijger M., Slijk E. (2006). Zwaaien in het onderwijs of onderwijs in het zwaaien (3), Lichamelijke opvoeding, 3: 30-32.

**Hoeboer J.**, Bosman M., Bakker J., Krijger M., Slijk E. (2006) Zwaaien in het onderwijs of onderwijs in het zwaaien (2), Lichamelijke opvoeding, 2006; 2: 29-30.

**Hoeboer J.**, Bosman M., Bakker J., Krijger M., Slijk E. (2006). Zwaaien in het onderwijs of onderwijs in het zwaaien (1), Lichamelijke opvoeding, 1: 27-29.

Bosman M., **Hoeboer J.** (2006). De Bouwstenen van het turnen op school (1e druk). GrafiSelect B.V., Haaksbergen.

**Hoeboer J.**, Bakker J. (2004). Wegwijs in het land van de overslag (2), Lichamelijke opvoeding, 11: 20-24.

**Hoeboer J.**, Bakker J. (2004). Wegwijs in het land van de overslag (1), Lichamelijke opvoeding, 10: 19-24.

**Hoeboer J.**, Bakker J. (2004). Turnen op school; met dank aan de turnsport (2), Lichamelijke opvoeding, 4: 36-40.

**Hoeboer J.**, Bakker J. (2004). Turnen op school; met dank aan de turnsport (1), Lichamelijke opvoeding, 2: 18-24.

Scientific conference proceedings

AIIESEP 2019 USA – oral presentation: **Hoeboer, J. J. A. A. M.**, Opstoel, K., Martelaer, K., Savelsbergh, G., & de Vries, S. I. Reproducibility, Validity and Feasibility of motor competence assessment instruments: a systematic review.

NASPSPA 2017 USA – oral presentation: **Hoeboer, J. J. A. A. M.**, Ongena, G., Krijger-Hombergen, M., Stolk, E., Savelsbergh, G. J. P., & de Vries, S. I. The Athletic Skills Track: age- and gender-related normative values of a motor skills test for 4- to 12-year-old children.

FIEP 2017 Luxembourg – oral presentation: Michiel Krijger, **Joris Hoeboer** en Sanne de Vries. Accuracy of PE teachers to estimate children's motor competence.

FIEP 2017 Luxembourg - **J. Hoeboer**, S.I. de Vries, G. Savelsbergh. What do children think of the Athletic Skills Track as a motor competence test in physical education?

FIEP 2017 Luxembourg - **J. Hoeboer**, M. Krijger, G. Ongena, E. Stolk, G. Savelsbergh, S.I. de Vries. The Athletic Skills Track: age- and gender-specific normative values of a motor skills test for 4- to 12-year-old children.



ACSM 2016 Australia – oral presentation: **Hoeboer, J. J. A. A. M.**, Ongena, G., Krijger-Hombergen, M., Stolk, E., Savelsbergh, G. J. P., & de Vries, S. I. The development of the Athletic Skills Track and the norm values for 6- to-12-year-old children.

ISBNPA 2016 South Africa – poster presentation: Hoeboer, J., Krijger-Hombergen, M., De Vries, S. Motor skill competence in relation to weight status and sports participation in 5-12 years old children.

ISBNPA 2016 South Africa – oral presentation: **Hoeboer, J.**, Krijger-Hombergen, M., Savelsbergh, G., & De Vries, S. Reliability and validity of an athletic skills track to assess motor competence among 4- to 12-year old children in physical education.

ECSS 2015 Sweden – oral presentation: **Hoeboer, J.**, De Vries, S., Krijger-Hombergen, M., Wormhoudt, R., Drent, A., Krabben, K., & Savelsbergh, G. Validity of the Athletic Skills Track to assess motor competence among 6- to 12-year old children in physical education.

## Other Media

NRO-praktijkonderzoek promotion movie (2018) (<https://www.nro-congres.nl/sessies/nro-verbindingsprijs-voor-onderzoekers/>)

Promotion movie 2basics (2018): <http://www.2basics.nl/portfolio-item/mq-schooltest/>

Article Telegraaf Dutch newspaper (21-12-2017) Kwart kinderen scoort onvoldoende op motorische test <https://www.telegraaf.nl/nieuws/1464258/kwart-kinderen-scoort-onvoldoende-op-motorische-test>

Based on the Telegraaf newspaper article: Dutch News Television (22-12-2017)

- <https://www.rtl.nl/video/d863a849-9366-33b9-80df-e832a0ae462c/> (10.00-11.49)
- <https://www.oudersvanu.nl/nieuws/kwart-nederlandse-kinderen-slechte-motoriek/>
- <https://www.rtlnieuws.nl/nederland/kwart-kinderen-heeft-slechte-motoriek-van-essentieel-belang-dat-er-nu-iets-verandert>
- [http://www.startpuntonderwijs.nl/nieuws\\_item/kinderen-klunziger/](http://www.startpuntonderwijs.nl/nieuws_item/kinderen-klunziger/)
- <http://www.amstelveenweb.com/nieuws-Kwart-van-de-Nederlandse-kinderen-zijn-slecht-mo&newsid=308689245>

Article NRC Dutch newspaper (08-10-2016), onderzoekskatern: Gewoon een koprol maken,

NRO-praktijkonderzoek promotion movie (2016) ([https://youtu.be/EkV\\_kV65xlc](https://youtu.be/EkV_kV65xlc)).

Awards

**Nomination NRO-VOR-praktijkprijs (2018)**

Nomination for applied research with highest impact in the professional working field

**Nomination NRO-VOR-praktijkprijs (2016)**

Nomination for applied research with highest impact in the professional working field

**Pim Breebaart Research Award (2015)**

Award for the best Applied Research Program; The Hague University of Applied Sciences

**Best Student Award (1999)**

Award for the best Bachelor student at The Hague University of Applied Sciences

## Appendix 1: Description of the Athletic Skills Tracks

Description of the items concerning the AST-1:

- (1) Alligator crawl; lie on the floor with your belly touching the mat and crawl staying low, using your arms and legs to move forward like an alligator;
- (2) Bunny hops over a bench; jump from side to side over the bench with hand support;
- (3) Travelling jumps; jump on both feet from one hoop to the next;
- (4) Throwing and catching a ball; throw the ball against the wall and catch it;
- (5) Kicking and stopping a ball; kick the ball against the wall and stop it;
- (6) Forward roll; bring the chin to the chest, use the hands to support the body, make floor contact with the back of the head, and roll progressively to the spine;
- (7) Backward roll; bring hands to ears as you bend your knees, begin rolling backwards in tucked position with the chin on the chest, place hands on the floor, and maintain in tucked position till the feet touch the floor;
- (8) Running backwards; run backwards as fast as possible around the cones;
- (9) Clambering; clamber over a vaulting box;
- (10) Jumping; jump over a string, land on both feet on the mat.

Description of the items concerning the AST-2:

- (1) Alligator crawl; lie on the floor with your belly touching the mat and crawl staying low, using your arms and legs to move forward like an alligator;
- (2) Walking forwards across a bench; keeping balance during walking forward on top of the bench;
- (3) Travelling jumps; jump on both feet from one hoop to the next;
- (4) Forward roll; bring the chin to the chest, use the hands to support the body, make floor contact with the back of the head, and roll progressively to the spine;
- (5) Backward roll; bring hands to ears as you bend your knees, begin rolling backwards in tucked position with the chin on the chest, place hands on the floor, and maintain in tucked position till the feet touch the floor;
- (6) Flank vault; place hands shoulder wide apart on top of the vaulting box, swing both legs from one side to the other side of the vaulting box;
- (7) Pencil roll; roll with your arms straight over your head (longitudinal);
- (8) Tumbling forward over a static trapeze, start in front support, bring chin to the chest, move upper body forward in tucked position while tumbling forward and holding the trapeze with both hands;
- (9) Hopscotch, jump on one leg around the cones;
- (10) Straight jump; start on vaulting box, jump into the trampoline and make a straight jump, land on both feet on the mat.

## Appendix 2: Test manual Athletic Skills Track

The test manual of the Athletic Skills Track (AST) has been developed as part of a project that was initiated by the VU University Amsterdam, ASM B.V. and The Hague University of Applied Sciences, The Netherlands.

In 2012, a pilot study was carried out by the Department of Human Movement Sciences, Faculty of Behavioural and Human Movement Science, VU University Amsterdam to assess children's motor skills while performing several tasks on an Athletic Skills Track. The results of the pilot study were promising and led to a larger research project conducted by The Hague University of Applied Sciences and the VU University Amsterdam to refine the track and examine its' validity and reliability. The results of this research project have shown that the AST offers PE teachers a feasible, valid and reliable tool to assess the motor skill level of children aged 4-12 years in a physical education (PE) setting (Hoeboer et al., 2016). In one PE lesson the motor skill level of 25-30 children can be assessed using an age-specific version of the AST. In this test manual, the three age-specific tracks that have been developed and validated for children aged 4-6 years (AST-1), 6-9 years (AST-2) and 9-12 years (AST-3) are explained in more detail, followed by a step-by-step plan for the teachers. Drawings and dimensions of the test and test forms are supplemented in the appendices.

Three age-specific Athletic Skills Tracks

Before the test can be conducted, it is important to read through this protocol and follow the step-by-step plan exactly. The test can be subdivided into three Athletic Skills Tracks (see also Table 1):

- AST-1 for children from groups 1 and 2 in The Netherlands (pre-school, kindergarten) in the age of 4- 6 years old.
- AST-2 for children from groups 3-5 (grades 1-3), in the age of 6-9 years old.
- AST-3 for children from groups 6-8 (grades 4-6), in the age of 9-12 years.

**Table 1:** Overview of AST per group and age.

	Group								Age								
	1	2	3	4	5	6	7	8	4	5	6	7	8	9	10	11	12
AST-1																	
AST-2																	
AST-3																	

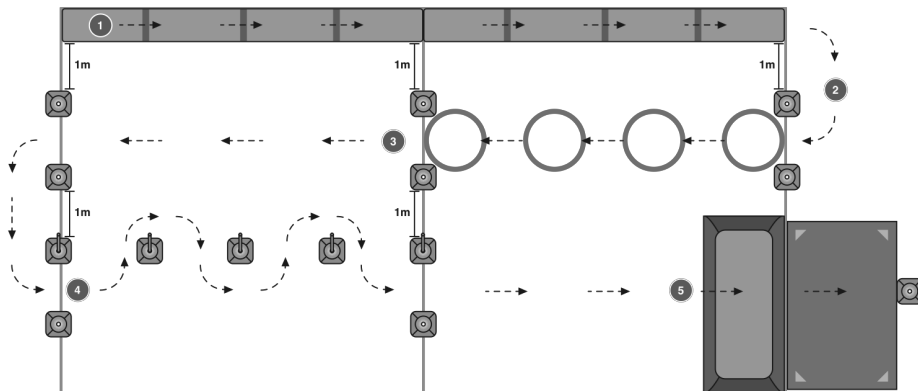
(AST-1 = Athletic Skills Track-1, AST-2 = Athletic Skills Track-2 and AST-3 = Athletic Skills Track-3)

The Athletic Skills Tracks have been developed in such a manner that it can be incorporated in a standard PE lesson, where activities are carried out in 3 sections of the gymnasium. While the test is conducted in one third of the gymnasium the other two sections can be used for the rest of the activities in the meantime. In the other sections, the children can practice independently on motor skills that have no influence on the test or test results.

When setting up the track, the drawings and dimensions of the test as shown in Appendix 1 should be followed precisely. As a variety of materials can be found in different inventories in a gymnasium, such as benches of various lengths, the drawings show the dimensions that must be fulfilled.

## Athletic Skills Track 1

AST-1 is suitable for children from groups 1 and 2 in The Netherlands (pre-school, kindergarten) in the age of 4 to 6 years old. In Figure 1 the map of AST-1 is shown. Further details about the dimensions can be found in Appendix 1. In Table 2 an overview of the tasks in AST-1 is presented. The number of the task corresponds to the number on the map in



**Figure 1.** AST-1 schematically displayed. (1) Walking, (2) Traveling jumps, (3) Alligator crawl, (4) Slaloming, and (5) Clambering

**Table 2:** Overview of the tasks in AST-1.

No.	Task	Arrangement	Dimensions
1	Walking (Balancing forwards) touching ribbons with 1 hand	2 Benches 3 Ribbons per bench	Length x width x height: 360 x 27 x 30 cm per bench
2	Traveling jumps (Walking forwards through the hoops)	4 Hoops	Diameter: 70 cm
3	Alligator Crawl (Walking forwards on all four limbs, belly facing the floor)	Cones on the ground	
4	Slaloming (Walking forwards)	5 Badminton posts or 5 Cones on the ground	
5	Clambering (Climbing over the box)	3 Box parts + top of the box 1 Mat 1 Cone	Length x width x height: 140 x 60 x 63 cm

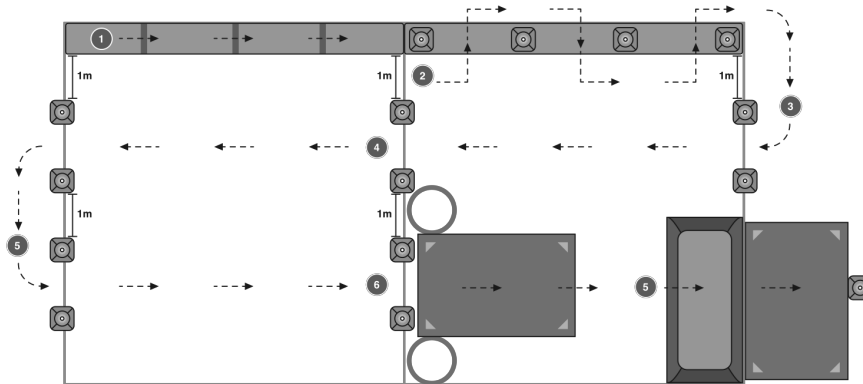
In AST-1, the following tasks are carried out in the following order:

1. Walking (Balancing forwards): The child starts on the bench. The teacher gives the starting signal and starts the stopwatch. When walking across the bench, the child touches the ribbons with one hand. The bench is in a normal position, seat facing up.
2. Traveling jumps (Walking forwards through the hoops): The child walks through the hoops. The order in which the feet are placed and the number of feet in the hoops are not important.
3. Alligator Crawl (Walking forwards on all four limbs): The child walks forward on all four limbs with the face pointing forwards and the belly facing the ground. The child starts in its position with both hands before the line of the two cones and may only stand up again when both hands pass the line of the following cones.
4. Slaloming (Walking forwards): The child slaloms forwards between the badminton posts (or cones).
5. Clambering (Climbing over the box): The child climbs over the box, lands on the mat and touches the cone. The teacher stops the stopwatch once the child has touched the cone.



## Athletic Skills Track 2

AST-2 is suitable for children from groups 3-5 in The Netherlands (primary school) in the age of 6 to 9 years old. In Figure 2 the map of AST-2 is shown. Further details about the dimensions of AST-2 can be found in Appendix 1. In Table 3 an overview of the tasks in AST-2 is presented. The number of the task corresponds to the number on the map in Figure 2.



**Figure 2.** AST-2 schematically displayed. (1) Walking, (2) Traveling jumps, (3) Hopscotch, (4) Alligator crawl (backwards), (5) Running (backwards), (6) Pencil roll, and (7) Clambering.

**Table 3.** Overview of the tasks in AST-2

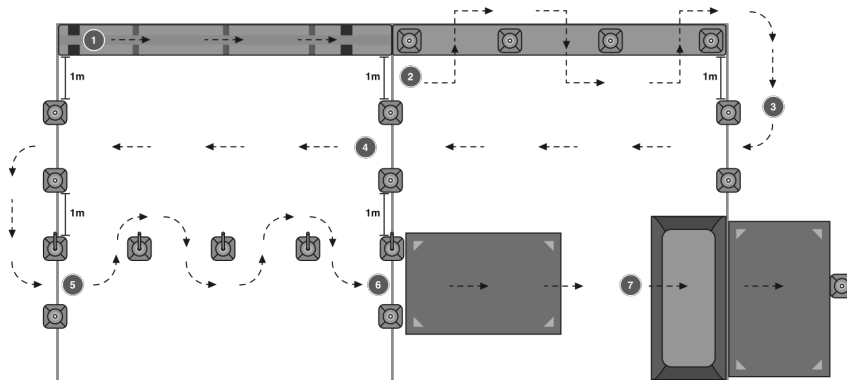
No.	Task	Arrangement	Dimensions
1	Walking (Balancing backwards, touching ribbons with 1 hand)	Bench	Length x width x height 360 x 27 x 30 cm
2	Traveling jumps (Vaulting forward)	Bench with 4 cones	Length x width x height 360 x 27 x 30 cm
3	Hopscotch (Hopping)	Cones on the ground	
4	Alligator crawl (Walking backwards on all four limbs, belly facing the floor)	Cones on the ground	
5	Running (Walking backwards)	Cones on the ground	
6	Pencil roll (Side roll)	1 Mat 2 Hoops	Length x width x height: 150 x 100 x 5 cm Diameter: 70 cm
7	Clambering (Climbing over the box)	4 Box parts + top of the box 1 Mat 1 Cone	Length x width x height: 140 x 60 x 76 cm

In AST-2, the following tasks are carried out in the following order:

1. Walking (Balancing backwards): The child begins on the bench. The teacher gives the starting signal and starts the stopwatch. When walking backwards across the bench, the child must touch the ribbons with one hand. The seat of the bench is facing upwards.
2. Traveling jumps (Vaulting forwards): The child performs three vaults over the bench between the cones. It does not matter whether the child starts at the left or right side of the bench.
3. Hopscotch (Hopping forwards): The child hops between the cones. The child can choose the foot.
4. Alligator crawl (Walking backwards on all four limbs): The child walks backwards on all four limbs with the face pointing towards the previous section of the track (hopping) and the belly facing the ceiling. The child starts in its position with both hands before the line of the two cones and may only stand up again when both hands pass the line of the following cones.
5. Running (Walking backwards): The child walks backwards between the cones.
6. Pencil roll (Side roll): The child places both hands in a hoop and both feet in the other hoop, then the child performs a side roll onto the mat (360°).
7. Clambering (Climbing over the box): The child climbs over the box, lands on the mat and touches the cone. The teacher stops the stopwatch once the child has touched the cone.

## Athletic Skills Track 3

AST-3 is suitable for children from groups 6, 7 and 8 in The Netherlands (primary school) in the age of 9 to 12 years old. In Figure 3 the map of AST-3 is shown. Further details about the dimensions of AST-3 can be found in Appendix 1. In Table 4 an overview of the tasks in AST-3 is presented. The number of the task corresponds to the number on the map in Figure 3.



**Figure 3.** AST-3 schematically displayed. (1) Walking (backwards), (2) Traveling jumps, (3) Hopscotch (backwards), (4) Alligator crawl (backwards), (5) Slaloming (backwards), (6) Forward roll, and (7) Clambering.

**Table 4.** Overview of the tasks in AST-3.

No.	Task	Arrangement	Dimensions
1	Walking (Balancing forwards, touching ribbons with 1 hand)	Upside down bench	Length x width x height: 360 x 7 x 30 cm
2	Traveling jumps (Vaulting forward)	Bench with 4 cones	
3	Hopscotch (Hopping backwards)	Cones on the ground	
4	Alligator crawl (Walking backward on all four limbs, belly facing the floor)	Cones on the ground	
5	Slaloming (Walking backwards)	5 Badminton posts or 5 Cones on the ground	
6	Forward roll	1 Mat	Length x width x height: 150 x 100 x 5 cm
7	Clambering (Climbing over the box)	5 Box parts and top of the box 1 Mat 1 cone	Length x width x height: 140 x 60 x 90 cm

In AST-3, the following tasks are carried out in the following order:

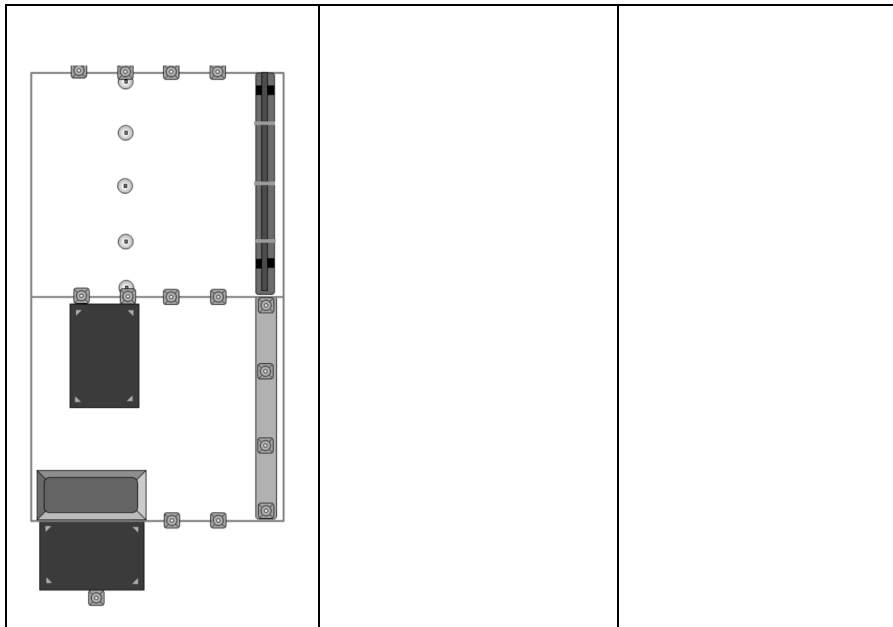
1. Walking (Balancing forwards): The child begins on the bench. The teacher gives the starting signal and starts the stopwatch. When walking forwards across the bench, the child must touch the ribbons with one hand. The bench is upturned; the seat is facing downwards.
2. Traveling jumps (Vaulting forwards): The child performs three vaults over the bench between the cones. It does not matter whether the child starts at the left or right side of the bench.
3. Hopscotch (Hopping backwards): The child hops backwards between the cones.
4. Alligator crawl (Walking backwards on all four limbs): The child walks backwards on all four limbs with the face pointing towards the previous section of the track (hopping) and the belly facing the ceiling. The child starts in its position with both hands before the line of the two cones and may only stand up again when both hands pass the line of the following cones.
5. Slaloming (Walking backwards): The child slaloms backwards between the badminton posts (or cones).
6. Forward roll: The child performs a forward roll.
7. Clambering (Climbing over the box): The child climbs over the box, lands on the mat and touches the cone. The teacher stops the stopwatch once the child has touched the cone

### Step-by-step plan for the teacher

In order to set up the Athletic Skills Tracks in the correct manner, a step-by-step plan has been drawn up for the user. It is important that you read through the step-by-step plan carefully before conducting the test.

### Preparation

The test can be carried out in one third of the gymnasium (see Figure 4). The other two sections of the gymnasium can be used for other activities. These activities must not influence the test or test results.



**Figure 4:** Organisation of the gymnasium into three sections.

### Setting up the track

The track must be set up before the children enter the gymnasium. As described, there are three different tracks, each for a different age group. The skills that the children must perform on each track are also different; take good note of this.

When setting up the track, the distances are vital. The distances are stated in the layout in Appendix 1. The distances are all standard dimensions, such as the width of a hoop or the length of a mat. These dimensions are derived from a standard inventory.

Conducting the test

### **Start of the lesson**

At the start of the lesson, a general overview of the structure of the lesson will be given. Ensure that children take into consideration the fact that they cannot disturb the area in which they are being tested. Apart from this, the lesson can be carried out as normal.

### **Demonstration by the test leader**

Once the children are ready to carry out the test, each group (8 to 10 children) is given a demonstration of the age-specific version of the test by the test leader (teacher).

### **Practice runs**

Once the demonstration has been carried out by the test leader, the children carry out 3 try-out trials. Within these trials, children practice the Athletic Skills Track with bare feet under the supervision of the test leader. During these practice runs, the test leader corrects any errors the children may be making. It is more efficient for the teacher to allow several children to practice on the track at the same time.

### **Test attempts**

One test attempt is carried out per child. In these test attempts, the children must complete the track without any errors and with bare feet. Only motor skill competence aspects must have an influence on the test, so if the track is misinterpreted, this is not measured. Only one child is allowed on the track at a time during the test. If a child makes an error, misses something or falls, then he/she may begin again, and the stopwatch is restarted.

### **Test results**

The test leader is responsible for recording the child's time to complete the track from start to end, by means of a stopwatch. The test leader makes sure the child takes on the starting position; standing on the bench at the start of the track. The test starts when the test leader gives a clear audible signal. Each test attempt begins with the child standing waiting on the bench at the start of the track. The test leader shouts: "ready, steady, go". On "go", he presses the stopwatch and the child must start the test. The test leader presses the stopwatch again once the child touches the cone at the end of the track. The time must be recorded in tenths of seconds, where 0.05 is rounded up to 0.1 seconds.

**Important notes**

- The track should be carried out with bare feet
- 3 practice runs
- 1 test attempt
- Only one child at a time on the track during the test
- Record the time to complete the track in tenths of seconds (for example 16.8 seconds)

**Note:**

In The Netherlands, all gymnasias have a standard inventory. This might not be the case in your country. For example, in The Netherlands a standard bench measures 3.6 metres. If such a bench is not available in your situation, the track should be adjusted in such a manner that the track remains within the stipulated square and the total distance that is covered does not change. For example, if only benches measuring 3 metres are available, then a distance of 1.20 metres must be left vacant between the first and second bench. After crossing the first bench, the child steps down from the bench and goes to the second bench.

If only longer benches than 3.6 metres are available, let the bench jut out somewhat and place a cone at the position where the children must start or make a transition to another part of the track

## Appendix 3: Descriptive statistics of respondents' characteristics

	Our sample	Population
<b>Gender</b>		
Boys	49.4%	51.3%
Girls	50.6%	48.7%
<b>Age</b>		
< 4	2.1%	9.0%
Age 5	4.7%	12.2%
Age 6	10.4%	12.3%
Age 7	15.9%	12.4%
Age 8	15.7%	12.1%
Age 9	14.7%	12.4%
Age 10	12.6%	12.5%
Age 11	13.1%	12.6%
Age 12	10.8%	4.5%

Descriptive statistics of the respondents in our sample as well as in the Dutch population (Centraal Bureau voor de Statistiek [CBS], 2016)



## Appendix 4: 10-item evaluation checklist

- 1) Sample characteristics (n, sex, age, weight, height, health status, ethnicity)**
  - $\geq 5$  sample characteristics are described (at least: n, sex, age, weight and height) (1)
  - 3-4 sample characteristics are described (0.5)
  - $\leq 2$  sample characteristics are described (0)
- 2) Protocol information (setting, individual/group assessment, duration, instructions, manual, test leader and training of the test leader)**
  - Complete information on protocol ( $\geq 5$  characteristics are described (at least: setting, individual/group, duration, manual, test leader) (1)
  - Some information on protocol (3-4 characteristics are described) (0.5)
  - Not clear at all ( $\leq 2$  characteristics are described) (0)
- 3) Content of the assessment (purpose, items, reference methods, outcome measures)**
  - Complete information on the content of the assessment (1)
  - Some information on the content of the assessment (0.5)
  - Very limited information on the content of the assessment (0)
- 4) Statistical analyses (tests, subgroup analysis), statistical software package and P value**
  - Complete information on statistical (1)
  - Some information on statistical analyses (0.5)
  - Very limited information on statistical analyses (0)
- 5) Reliability, internal consistency (Cronbach's Alfa or Blend and Altman Plots)**
  - Present (1)
  - Absent (0)
- 6) Reliability, intra-rater / test-retest reliability (ICC of Pearson)**
  - Present (1)
  - Absent (0)
- 7) Betrouwbaarheid, inter-rater betrouwbaarheid (ICC of Pearson)**
  - Present (1)
  - Absent (0)
- 8) Validity, Construct (confirmatory factor analysis)**
  - Present (1)
  - Absent (0)
- 9) Validity, Content (exploratory factor analysis)**
  - Present (1)
  - Absent (0)
- 10) Concurrent /criterion validity (correlation)**
  - Present (1)
  - Absent (0)

## Appendix 5: Reference list for included studies

- 1) Kita, Y., Suzuki, K., Hirata, S., Sakihara, K., Inagaki, M., & Nakai, A. (2016). Applicability of the Movement Assessment Battery for Children-to Japanese children: A study of the Age Band 2. *Brain and Development*, 38(8), 706-713.
- 2) Ré, Alessandro HN, et al. "Comparison of motor competence levels on two assessments across childhood." *Journal of sports sciences* 36.1 (2018): 1-6.
- 3) Kambas, A., & 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(1), 65-72.
- 4) Lane, H., & Brown, T. (2015). Convergent validity of two motor skill tests used to assess school-age children. *Scandinavian journal of occupational therapy*, 22(3), 161-172.
- 5) Psotta, R., & Brom, O. (2016). Factorial structure of the movement assessment battery for children test—second edition in preschool children. *Perceptual and motor skills*, 123(3), 702-716.
- 6) Vandorpe, B., Vandendriessche, J., Lefèvre, J., Pion, J., Vaeyens, R., Matthys, S. A., & Lenoir, M. (2011). The KörperkoordinationsTest für Kinder: reference values and suitability for 6–12-year-old children in Flanders. *Scandinavian journal of medicine & science in sports*, 21(3), 378-388.
- 7) Kim, S., Kim, M. J., Valentini, N. C., & Clark, J. E. (2014). Validity and reliability of the TGMD-2 for South Korean children. *Journal of Motor Behavior*, 46(5), 351-356.
- 8) Psotta, R., & Abdollahipour, R. (2017). Factorial Validity of the Movement Assessment Battery for Children—2nd Edition (mabc-2) in 7-16-year-olds. *Perceptual and motor skills*, 124(6), 1051-1068.
- 9) Hoeboer, J., De Vries, S., Krijger-Hombergen, M., Wormhoudt, R., Drent, A., Krabben, K., & Savelsbergh, G. (2016). Validity of an Athletic Skills Track among 6-to 12-year-old children. *Journal of sports sciences*, 34(21), 2095-2105.
- 10) Lander, N., Morgan, P. J., Salmon, J., Logan, S. W., & Barnett, L. M. (2017). The reliability and validity of an authentic motor skill assessment tool for early adolescent girls in an Australian school setting. *Journal of science and medicine in sport*, 20(6), 590-594.
- 11) Herrmann, C., & Seelig, H. (2017). Structure and Profiles of Basic Motor Competencies in the Third Grade—Validation of the Test Instrument MOBAK-3. *Perceptual and motor skills*, 124(1), 5-20.

- 12) Issartel, J., McGrane, B., Fletcher, R., O'Brien, W., Powell, D., & Belton, S. (2017). A cross-validation study of the TGMD-2: The case of an adolescent population. *Journal of science and medicine in sport, 20*(5), 475-479.
- 13) Aye, T., Oo, K. S., Khin, M. T., Kuramoto-Ahuja, T., & Maruyama, H. (2017). Reliability of the test of gross motor development second edition (TGMD-2) for Kindergarten children in Myanmar. *Journal of physical therapy science, 29*(10), 1726-1731.
- 14) Hoeboer, J., Krijger-Hombergen, M., Savelsbergh, G., & De Vries, S. (2018). Reliability and concurrent validity of a motor skill competence test among 4-to 12-year old children. *Journal of sports sciences, 36*(14), 1607-1613.
- 15) Kakebeeke, T. H., Knaier, E., Köchli, S., Chaouch, A., Rousson, V., Kriemler, S., & Jenni, O. G. (2016). Comparison between the Movement ABC-2 and the Zurich Neuromotor Assessment in preschool children. *Perceptual and motor skills, 123*(3), 687-701.
- 16) Kim, S., Kim, M. J., Valentini, N. C., & Clark, J. E. (2014). Validity and reliability of the TGMD-2 for South Korean children. *Journal of Motor Behavior, 46*(5), 351-356.
- 17) Kambas, A., & Venetsanou, F. (2014). The Democritos Movement Screening Tool for preschool children (DEMOST-PRE©): Development and factorial validity. *Research in developmental disabilities, 35*(7), 1528-1533.
- 18) Fransen, J., D'Hondt, E., Bourgois, J., Vaeyens, R., Philippaerts, R. M., & Lenoir, M. (2014). Motor competence assessment in children: Convergent and discriminant validity between the BOT-2 Short Form and KTK testing batteries. *Research in developmental disabilities, 35*(6), 1375-1383.
- 19) Kim, C. I., Han, D. W., & Park, I. H. (2014). Reliability and validity of the Test of Gross Motor Development-II in Korean preschool children: applying AHP. *Research in developmental disabilities, 35*(4), 800-807.
- 20) Valentini, N. C., Ramalho, M. H., & Oliveira, M. A. (2014). Movement Assessment Battery for Children-2: Translation, reliability, and validity for Brazilian children. *Research in Developmental Disabilities, 35*(3), 733-740.
- 21) Barnett, L. M., Minto, C., Lander, N., & Hardy, L. L. (2014). Interrater reliability assessment using the Test of Gross Motor Development-2. *Journal of Science and Medicine in Sport, 17*(6), 667-670.
- 22) Holm, I., Tveter, A. T., Aulie, V. S., & Stuge, B. (2013). High intra-and inter-rater chance variation of the movement assessment battery for children 2, ageband 2. *Research in developmental disabilities, 34*(2), 795-800.

- 23) Hua, J., Gu, G., Meng, W., & Wu, Z. (2013). Age band 1 of the Movement Assessment Battery for Children-: Exploring its usefulness in mainland China. *Research in Developmental Disabilities, 34*(2), 801-808.
- 24) Valentini, N. C. (2012). Validity and reliability of the TGMD-2 for Brazilian children. *Journal of motor behavior, 44*(4), 275-280.
- 25) Kambas, A., Venetsanou, F., Giannakidou, D., Fatouros, I. G., Avloniti, A., Chatzinikolaou, A., ... & Zimmer, R. (2012). The Motor-Proficiency-Test for children between 4 and 6 years of age (MOT 4–6): An investigation of its suitability in Greece. *Research in Developmental Disabilities, 33*(5), 1626-1632.
- 26) Logan, S. W., Robinson, L. E., & Getchell, N. (2011). The comparison of performances of preschool children on two motor assessments. *Perceptual and motor skills, 113*(3), 715-723.
- 27) Smits-Engelsman, B. C., Niemeijer, A. S., & van Waelvelde, H. (2011). Is the Movement Assessment Battery for Children-a reliable instrument to measure motor performance in 3 year old children? *Research in developmental disabilities, 32*(4), 1370-1377.
- 28) Ellinoudis, T., Evaggelinou, C., Kourtessis, T., Konstantinidou, Z., Venetsanou, F., & Kambas, A. (2011). Reliability and validity of age band 1 of the Movement Assessment Battery for Children–Second Edition. *Research in Developmental Disabilities, 32*(3), 1046-1051.
- 29) Zuvela, F., Bozanic, A., & Miletic, D. (2011). POLYGON-A new fundamental movement skills test for 8 year old children: Construction and validation. *Journal of sports science & medicine, 10*(1), 157.
- 30) Wagner, M. O., Kastner, J., Petermann, F., & Bös, K. (2011). Factorial validity of the Movement Assessment Battery for Children-2 (age band 2). *Research in developmental disabilities, 32*(2), 674-680.
- 31) Cools, W., De Martelaer, K., Vandaele, B., Samaey, C., & Andries, C. (2010). Assessment of movement skill performance in preschool children: Convergent validity between MOT 4-6 and M-ABC. *Journal of sports science & medicine, 9*(4), 597.
- 32) Engel-Yeger, B., Rosenblum, S., & Josman, N. (2010). Movement Assessment Battery for Children (M-ABC): establishing construct validity for Israeli children. *Research in Developmental Disabilities, 31*(1), 87-96.
- 33) Smits-Engelsman, B. C., Fiers, M. J., Henderson, S. E., & Henderson, L. (2008). Interrater reliability of the movement assessment battery for children. *Physical therapy, 88*(2), 286-294.

- 34) Croce, R. V., Horvat, M., & McCarthy, E. (2001). Reliability and concurrent validity of the movement assessment battery for children. *Perceptual and motor skills*, 93(1), 275-280.
- 35) Wong, K. Y. A., & Yin Cheung, S. (2010). Confirmatory factor analysis of the Test of Gross Motor Development-2. *Measurement in Physical Education and Exercise Science*, 14(3), 202-209.
- 36) Herrmann, C., Gerlach, E., & Seelig, H. (2015). Development and validation of a test instrument for the assessment of basic motor competencies in primary school. *Measurement in Physical Education and Exercise Science*, 19(2), 80-90.
- 37) Chow, S. M., & Henderson, S. E. (2003). Interrater and test-retest reliability of the Movement Assessment Battery for Chinese preschool children. *American Journal of Occupational Therapy*, 57(5), 574-577.
- 38) Rousson, V., Gasser, T., Caflisch, J., & Largo, R. (2008). Reliability of the Zurich neuromotor assessment. *The Clinical Neuropsychologist*, 22(1), 60-72.

Appendix 6: Table reliability included studies

General info		Reliability				Quality of the study (0-9)	Article (or)	
Intervention	Author(s)	Year	n	Age	Test-retest Intra-rater Inter-rater Cb.4			Intra-rater ICC/ kappa
9	AST	Hoebner et al., 2016	460 children (252 boys and 217 girls)	6-12	4.0	-	-	9
14	AST		500 children (452 boys and 448 girls)	4-12	3.5	Internal consistency of the AST as represented by Cronbach's $\alpha$ was at or above the acceptable level of 0.70 (Cronbach, 1951), i.e. for AST-1: $\alpha = 0.784$ ; for AST-2: $\alpha = 0.700$ and for AST-3: $\alpha = 0.746$ .	intra-rater coefficients between the first and second trial were 0.851 (95% CI 0.826-0.876) for AST-1 and 0.891 (95% CI 0.870-0.908) for AST-2	6.5
3	BOT-2 (short form)		377 (191 boys, 186 girls)	5-8.3	1.0	Cronbach alpha for all the items was 0.97.	-	3.0
18	BOT-2 (short form)		3488 children (1300 boys and 1188 girls)	6-11	2.0	-	-	3.0
4	BOT-2		30 (25 boys, 25 girls)	7-16	3.5	-	-	4.5
10	GAMSA		34 children (0 boys and 34 girls)	Mean 12.6	3.0	-	-	5.0
17	DEMOISE-FRE		453 children (258 boys and 197 girls)	4-6	2.0	-	test-retest reliability (i.e., the aggregate of skill and time stability) for the DEMOISE-FRE was 0.88 (95% CI = 0.86-0.90) and isolated skill score (delta ICC = 0.85)	4.0
29	FANS-POLYGOON		95 children (48 boys and 47 girls)	8	2.0	-	ICC for the FANS-POLYGOON showed a very high result (0.98) and confirmed the test's reliability	4.0
18	KTK		3488 children (1300 boys and 1188 girls)	6-11	2.0	-	-	3
2	KTK		424 children (224 boys and 200 girls)	5-10	3.5	-	-	4.5
6	KTK		2070 children (1297 boys and 1172 girls)	6-12	3.5	-	-	4.5
5	M-ABC-2		183 children (106 boys and 197 girls)	3-6	2.0	-	-	3.0
30	M-ABC-2		323 children (169 boys and 154 girls)	7-10	2.0	-	-	3.0
31	M-ABC-2		48 children (25 boys and 23 girls)	5-6	3.0	-	-	3.0
32	M-ABC-2		249 children (209 boys and 236 girls)	4-12	2.0	-	-	3.0
8	M-ABC-2		484 children (240 boys and 244 girls)	7-16	2.5	-	-	3.5
26	M-ABC-2		15 boys (15 boys and 17 girls)	3-6	2.5	-	-	3.5
33	M-ABC-2		131 nurses (none children)	4-12	3.0	-	Kappa coefficients ranged from 0.95-1.00	4.0
34	M-ABC-2		105 children (67 boys and 39 girls)	5-12	2.0	-	The ICC values ranged from 0.92 to 0.98	4.0
37	M-ABC-2		138 children (76 boys and 62 girls)	4-6	2.0	-	The ICC values overall of 0.86	4.0
4	M-ABC-2		25 children (25 boys, 25 girls)	7-16	3.5	-	-	4.5
22	M-ABC-2		45 children (22 boys and 23 girls)	8-9	3.0	-	The ICC values ranged from 0.23 to 0.76	5.0
1	M-ABC-2		132 children (84 boys and 48 girls)	7.0-10.8	3.5	Cronbach's alpha coefficient for the eight test items was 0.602	-	5.5
27	M-ABC-2		108 children (75 boys and 23 girls)	7-4	3.0	Cronbach's $\alpha$ ranges from 0.70 to 0.97	-	6.0
28	M-ABC-2		183 children (98 boys and 49 girls)	3-5	3.0	Cronbach's $\alpha$ ranges from 0.51 to 0.70	-	6.0
20	M-ABC-2		844 children (440 boys and 404 girls)	3-13	3.5	the 3 subscales ( $\alpha = 0.78$ ), subtest (nominal decrement: $\alpha = 0.77$ ), ball skills ( $\alpha = 0.52$ , balance: $\alpha = 0.77$ )	ICC values ranged from 0.61 to 0.90	8.5
25	M-ABC-2		103 children (61 boys and 48 girls)	3-6	3.0	Cronbach's alpha value of all eight items together was 0.502	ICC values ranged from 0.68 to 0.85	9.0
11	MORAK		823 children (465 boys and 358 girls)	8-10	2.5	-	The ICC values ranged from 0.892 to 0.939	3.5
36	MORAK		107 children (44 boys and 173 girls)	6-7	3.0	-	-	4.0

General info			Reliability				Quality of the Article (0-9)	Article (nr)				
nr	Instrument	Author(s)	year	n	Age	Pre-ocul info (score)			Internal consistency	Intra-Rater Reliability ICC-Chaparr	Test-retest Reliability ICC	Inter-rater Reliability ICC-Chaparr
31	MOT 4-6			48 children (25 boys and 23 girls)	3-6	3.0	-	-	-	-	3.0	31
25	MOT 4-6			32 children (15 boys and 17 girls)	4-6	3.0	Cronbach's alpha coefficient value was 0.901	ICC for test-retest criterion was found to be 0.87	-	-	5.0	25
26	TGMD-2			32 children (15 boys and 17 girls)	3-6	2.5	-	-	-	-	3.5	26
12	TGMD-2			844 children (456 boys and 388 girls)	11-12	3.0	-	-	-	-	4.0	12
21	TGMD-2			107 children (56 boys and 51 girls)	4-8	3.0	-	-	-	-	4.0	21
35	TGMD-2			650 children (332 boys and 318 girls)	3-10	3.0	-	-	-	-	4.0	35
2	TGMD-2			424 children (224 boys and 200 girls)	3-10	3.5	-	-	-	-	4.5	2
19	TGMD-2			121 children (71 boys and 50 girls)	5-6	3.0	Cronbach's alpha coefficient of pre-weighted TGMD-II scores was 0.72. The Cronbach's alpha of post-weighted TGMD-II scores was 0.73. Cronbach's alpha coefficients for locomotor skills and object control skills were 0.53 and 0.68, respectively. The alpha values for post-weighted scores of the locomotor skills were 0.53, for locomotor skill and object control skill, respectively.	The results of the test-retest reliability prevented acceptable reliability for ICC and strong positive to very strong positive degrees of correlation for Pearson coefficient between Day 1 and Day 2 assessment.	-	-	5.0	19
13	TGMD-2			50 children (23 boys and 27 girls)	4-6	3.5	-	The test-retest reliability results of the Pearson's test indicated a very strong positive to strong positive correlation between the raw scores of the locomotor ( $r = 0.90, p < .0001$ ) and the object control subscales ( $r = 0.88, p < .0001$ )	All the values of Cronbach's alpha and ICC coefficients for all the measures in the inter-rater reliability	-	5.5	13
7	TGMD-2			141 children	3-10	2.6	Cronbach's alpha coefficient analyses showed a very high internal consistency for each of the subscales as well as for each subset independently (locomotor, Cronbach's alpha = 0.82; object control, Cronbach's alpha = 0.73)	The test-retest reliability results of the Pearson's test indicated a very strong positive to strong positive correlation between the raw scores of the locomotor ( $r = 0.90, p < .0001$ ) and the object control subscales ( $r = 0.85, p < .0001$ )	The inter-rater reliability among the three evaluators for locomotor and object control scores as well as the TGMD-2 total score	The inter-rater reliability among the three evaluators for locomotor and object control scores as well as the TGMD-2 total score	6.0	7
16	TGMD-2			141 children	3-10	1.5	The Cronbach's alpha coefficient analyses showed a very high internal consistency for each of the subscales as well as for each subset independently (locomotor, Cronbach's alpha = 0.82; object control, Cronbach's alpha = 0.73)	The test-retest reliability results of the Pearson's test indicated a very strong positive to strong positive correlation between the raw scores of the locomotor ( $r = 0.90, p < .0001$ ) and the object control subscales ( $r = 0.85, p < .0001$ )	The inter-rater reliability among the three evaluators for locomotor and object control scores as well as the TGMD-2 total score	The inter-rater reliability among the three evaluators for locomotor and object control scores as well as the TGMD-2 total score	6.5	16
24	TGMD-2			2674 children (1352 boys and 1322 girls)	3-10	3.0	-	The Pearson test-retest reliability results showed a strong positive and significant correlation ( $r = .90, p < .0001$ ) for locomotor and object control (ICC = 0.90) and for the test-retest correlations were observed for all locomotor and object control subscales. The intratester reliability coefficients ( $\alpha = .92-.99$ ) indicated strong and consistent concordance among the	The inter-rater reliability for the locomotor ( $\alpha = .89$ ) and object control ( $\alpha = .89$ ) subscales was also high. Significant and positive test-retest correlations were observed for all locomotor and object control subscales. The intratester reliability coefficients ( $\alpha = .87-.92$ ) indicated strong and consistent concordance among the	The inter-rater reliability for the locomotor ( $\alpha = .89$ ) and object control ( $\alpha = .89$ ) subscales was also high. Significant and positive test-retest correlations were observed for all locomotor and object control subscales. The intratester reliability coefficients ( $\alpha = .87-.92$ ) indicated strong and consistent concordance among the	7.0	24
15	ZNA			47 children (26 boys and 21 girls)	3-5	2.5	-	ICC ranged from 0.90-0.90	-	-	3.5	15
38	ZNA			102 children (52 boys and 50 girls)	6-12	1.5	-	ICC ranged from 0.65-0.80	-	-	3.5	38

Appendix 7: Table validity included studies

General info		Validity				Quality of the study (0-9)	Article (n)				
nr	Instrument	Author(s)	year	n	Age			Prevalence of imbalance (score)	Reference method	Construct validity factor analyses	Concurrent Correlation (r)
9	AST			463 children (252 boys and 211 girls)	6-12	4.0	KTK	-	The correlation between AST-1 and KTK was (girls: $r = 0.50$ , $P < 0.01$ ; boys: $r = 0.53$ , $P < 0.01$ ). For AST-2 the correlation between AST-2 and the KTK was (girls: $r = 0.44$ , $P < 0.01$ ; boys: $r = 0.48$ , $P < 0.01$ ). The concurrent validity of the AST in AST-1: $r = -0.747$ , $p < 0.01$ ; AST-2: $r = -0.64$ , $p < 0.01$ ; and AST-3: $r = -0.602$ , $p < 0.01$ .	0.0	9
14	AST			910 children (482 boys and 428 girls)	4-12	3.5	KTK	-	The concurrent validity of the AST in AST-1: $r = -0.747$ , $p < 0.01$ ; AST-2: $r = -0.64$ , $p < 0.01$ ; and AST-3: $r = -0.602$ , $p < 0.01$ .	6.5	14
3	BOT-2 (short form)			577 (191 boys, 186 girls)	5-8	1.0	-	three factors had eigenvalues greater than 1.0 and accounted for 54.14% of the score variance: Gross and Fine Motor Proficiency (3.19%), Eye-Hand Coordination (4.5%), Visual-motor Control (7.2%).	-	3.0	3
18	BOT-2 (short form)			90 (25 boys, 25 girls)	7-16	3.5	M-ABC	-	significant correlations ( $0.60 \geq r \leq 0.64$ , $p < 0.001$ ) between BOT-2 Short Form total and gross motor composite scores and KTK were also found in each age group separately for the correlation between BOT-2 Short Form total and KTK ( $r = 0.60$ , $p < 0.001$ ). KTK Motor Quotient, significant correlation coefficients of 0.30 $\geq r \leq 0.37$ ( $p < 0.001$ ) were found for each age group.	3.0	18
4	BOT-2			90 (25 boys, 25 girls)	7-16	3.5	M-ABC	-	a positive strong correlation was found between the BOT-2 total Motor Composite and the ABB-MABC-2 Total Test Score (rho = 0.60, $p < 0.01$ )	4.5	4
10	CAMSA			34 children (0 boys and 34 girls)	Mean 12.6	3.0	Victorian Assessment	-	Spearman's Rho rank-order analysis using a non-rated test of the relationship between the scores on the CAMSA using their final finishing position of students in the CAMSA using their final CAMSA score and Victorian PMS Assessment in Test 1 (rs = 0.60, $p < 0.001$ ). The correlation between the scores on the KTK Motor Composite with the total skill score of the Victorian PMS Assessment, the correlation was slightly weaker, but still considered strong (rs = 0.60, $p < 0.05$ )	5.0	10
17	DEMOST-PRE			435 children (238 boys and 197 girls)	4-6	2.0	-	Kaiser-Meyer-Olkin index of .85 yielded two components (eigenvalues of 1.89 and 1.19) which explained 33.2% of the variance. The first component explained 33.2% of the variance and was named 'gross motor' and included the following activities: arm and hand carrying and placing a ball into a box, overhead walking in backward direction, stepping through 3 vertical bars, and carrying a ball into a box. The second component explained 20.9% of the variance and encompassed all other items that were implemented in the factor analysis (tapping, overhead walking, stepping through 3 vertical bars, picking up coins and placing them in a box, exchanging beads).	-	4.0	17
29	FMS-POLYGON			95 children (48 boys and 47 girls)	8	2.0	TGM2-2	-	Correlation analysis between the newly constructed FMS-POLYGON and TGM2-2 revealed a significant correlation ( $r = 0.60$ , $p < 0.05$ ) which indicates a high positive relationship between performance time in FMS-POLYGON test means better result.	4.0	29
18	KTK			2485 children (1300 boys and 1185 girls)	6-11	2	BOT-2	-	unlike TGM2-2, BOT-2 Short Form and KTK were found between total BOT-2 Short Form score ( $r = 0.61$ , $p < 0.001$ ) and KTK Motor Quotient and BOT-2 Short Form gross motor score ( $r = 0.59$ , $p < 0.001$ ). A weaker but significant correlation ( $r = 0.25$ , $p < 0.001$ ) was found between BOT-2 Short Form fine motor composite score and KTK Motor Quotient. Significant correlations ( $0.60 \geq r \leq 0.64$ , $p < 0.001$ ) between BOT-2 Short Form total and gross motor composite scores and KTK Motor Quotient, significant correlation coefficients of 0.30 $\geq r \leq 0.37$ ( $p < 0.001$ ) were found for each age group.	3.0	18
2	KTK			424 children (224 boys, 200 girls)	5-10	3.5	TGM2-2	-	All the correlations between the total raw scores on the TGM2-2 and KTK were statistically significant in all age groups, but the correlation between the total raw scores on the TGM2-2 and KTK was lowest at 6-10 years old ( $r = 0.38$ , $p < 0.22$ ) to moderate at 7-8 ( $r = 0.50$ , $p < 0.25$ ) and 5-6 years old ( $r = 0.52$ , $p < 0.27$ ).	4.5	2
6	KTK			509 children (297 boys and 112 girls)	6-12	3.5	KTK	-	the correlation between the scores on the KTK for the German sample ( $n = 143$ ) was statistically significant ( $r = 0.60$ , $p < 0.001$ )	4.5	6







General info		Validity										
nr	Instrument	Author(s)	Year	n	Age	Pre-Cell into (Cron's)	Reference method	Construct Construct validity factor analysis	Content Factor Analysis	Concurrent Correlation (r)	Quality of the study (0 - 9)	Article (nr)
24	TOMD-2			2574 children (1322 girls) and 1132 girls)	3 - 10	3.0	M-ABC-2	The results of the RMSEA (.05), 90% confidence interval [0.06, 0.04], and CFI (.98), and AIC (0.33) provided support for the two-factor model.	-	Person correlations indicated a significant positive, and small to moderate correlation between TOMD-2 and M-ABC explained only 7.29% of the variance in M-ABC. The correlations between TOMD-2 and M-ABC were moderate to large for the 3-4 years old revealed that for the young children, at ages 4 (r = .42, p = .05) and 5 years old (r = .56, p = .002), the associations were moderate to large. In contrast, for the 6 to 10 years old the correlations were small and nonsignificant (r = .14-.30, p > .05).	7.0	24
15	ZVA			47 children (26 boys and 21 girls)	3 - 5	2.5	M-ABC-2	-	-	The correlation between the summary scores of the two tests was 77.95% (CI = 0.82, 0.87) and increased to 0.84 (95% CI = 0.73, 0.91) when considering the individual components. When considering individual components, the largest correlations were observed between the verbal score for the motor subscale (r = 0.82, 0.73) and ZMAS-5 (r = .59, 0.56, 0.54, 0.52, 0.51) and the motor subscale (r = 0.82, 0.73) and M-ABC-2 and static balance in ZMAS-5 (r = 0.68, 0.57, 0.46, 0.41) and between balance in M-ABC-2 and the dynamic balance in ZMAS-5 (r = 0.82, 0.73) (CI = 0.74, 0.77).	3.5	15
38	ZVA			30 children (12 boys and 18 girls)	6 - 12	1.5	-	-	-	-	3.5	38

Appendix 8: Evaluation of all included studies

Level of evidence for the reliability and validity of MC measurement tools per study									
		Reliability			Validity		Quality		Studies
Instrument		Internal consistency Cb α (< 0.60 = -, 0.60 - 0.70 = +/-, > 0.70 = +)	Intra-Rater or Test-retest Reliability ICC/kappa/r (ICC (< 0.50 = -, 0.50-0.80 = +/-, > 0.80 = +))	Inter-rater Reliability (ICC (< 0.40 = -, 0.40-0.80 = +/-, > 0.80 = +))	Construct (Values < 0.50 = -, 0.50 - 0.70 = +/-, > 0.70 = +)	Concurrent R (< 0.50 = -, 0.50 - 0.70 = +/-, > 0.70 = +)	Score < 4 = -, 4 - 6 = +/-, ≥ 6 = +		Number
9	AST		+			+	+	6.0	9
14	AST	+	+			+	+	6.5	14
	AST	++	++			++	+	6.3	
4	BOT-2					+	±	4.5	4
3	BOT-2 (short form)	+			±		-	3.0	3
18	BOT-2 (short form)					-	-	3.0	18
	BOT	+			±	±	-	3.5	
10	CAMSA		+			±	±	5.0	10
	CAMSA	+	+			±	±	5.0	
17	DEMOST-PRE		+				±	4.0	17
	DEMOST-PRE		+				±	4.0	
29	FMS-POLYGOON		+			+	±	4.0	29
	FMS-POLYGOON		+			+	±	4.0	
2	KTK					-	±	4.5	2
6	KTK					±	±	4.5	6
18	KTK					±	-	3.0	18
	KTK					±	±	4.0	
1	M-ABC-2	±			+		±	5.5	1
4	M-ABC-2					+	±	4.5	4
5	M-ABC-2				+		-	3.0	5
8	M-ABC-2				+		-	3.5	8
20	M-ABC-2	±	±	+		-	+	8.5	20
22	M-ABC-2		-	-			±	5.0	22
23	M-ABC-2	-	+	+	-	±	+	9.0	23
26	M-ABC-2					-	-	3.5	26
27	M-ABC-2	+	+	±			+	6.0	27
28	M-ABC-2	-	±		±		+	6.0	28
30	M-ABC-2				+		-	3.0	30
31	M-ABC-2						-	3.0	31
32	M-ABC-2						-	2.0	32
33	M-ABC-2			+			±	4.0	33
34	M-ABC-2		+			±	±	4.0	34
37	M-ABC-2		±	±			±	4.0	37

Level of evidence for the reliability and validity of MC measurement tools per study							
Instrument	Reliability			Validity		Quality	Studies
	Internal consistency Cb $\alpha$ (< 0.60 = -, 0.60 - 0.70 = +/-, > 0.70 = +)	Intra-Rater or Test-retest Reliability ICC/kappa <sup>2</sup> (ICC (< 0.50 = -, 0.50-0.80 = +/-, > 0.80 = +)	Inter-rater Reliability (ICC (< 0.40 = -, 0.40-0.80 = +/-, > 0.80 = +)	Construct (Values < 0.50 = -, 0.50-0.70 = +/-, > 0.70 = +)	Concurrent R (< 0.50 = -, 0.50-0.70 = +/-, > 0.70 = +)	Score < 4 = -, 4 - 6 = +/-, $\geq 6$ = +	Number
	M-ABC-2	+++	+++	+++	+++	±	4.7
11	MOBAK			+		-	3.5
36	MOBAK			+		±	4.0
	MOBAK			+		-	3.5
25	MOT 4-6	+	+			+/-	5.0
31	MOT 4-6					-	3.0
	MOT 4-6	+	+			±	4.0
2	TGMD-2		+			-	4.5
7	TGMD-2	+		±	+	+	6.0
12	TGMD-2					±	4.0
13	TGMD-2		+	+		±	5.5
19	TGMD-2	+/-				-	6.5
21	TGMD-2			±		±	5.0
24	TGMD-2		+	±	+	-	4.0
26	TGMD-2					-	7.0
35	TGMD-2				+	±	4.0
	TGMD-2	+++	+++	+++	+++	+	5.3
15	ZNA					+	3.5
38	ZNA		±	±		-	3.5
	ZNA		±	±	+	-	3.5

First character: “-“= poor reliability or validity, “+/-“ = moderate or mixed reliability or validity, “+“ = good reliability or validity.

Second character = “-“< 4 points for the quality of the study, “±” 4 – 6 points for the quality of the study, “+“ > 6 points for the quality of the study)

Third extra character = if there are 5 or more good quality studies



This thesis describes the development of the Athletic Skills Track. The aim of this thesis is to examine the reliability, validity and feasibility of this new MC assessment tool: The Athletic Skills Track to assess fundamental movement skills among 4- to 12-year-old children in a physical education setting.

There seems to be an urgency to increase our understanding of motor skill development that is also recognised by PE teachers. They are willing to monitor motor skill competence of children more objectively. Unfortunately, there seem to be some practical shortcomings in the available assessment tools. It takes at least 20 min to measure one individual child. Furthermore, special test materials and extensive knowledge of the test protocols are required to be able to conduct the tests. Those shortcomings seem to be the reason many PE teachers currently do not use motor skill competence tests structurally. Therefore, a reliable and valid assessment tool that fits the PE setting is needed.