8
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.
Validly 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 Rodgerson (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?
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 de 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 is 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


