Gender differences in motor skill performance on the PERF-FIT

Alexandra Knauerhase

(s3661253)

University of Groningen

E-mail: a.s.knauerhase@rug.nl

Date: 07-02-2022

Author Note

This article was written as a Bachelor's thesis under supervision of Dr. Dorothee Jelsma.

Abstract

The present study investigated the gender differences in motor performance by means of a cross-sectional field study including 149 school-aged children from the Netherlands and Brazil, who completed the Performance and Fitness Battery (PERF-FIT). Boys and girls were compared based on their raw scores on the single test items. A significant difference was found for the item *dynamic balance,* with girls scoring higher than boys. The differences on all other items were non-significant and overall, effect sizes were small. Although the utility of gender-specific norms for the PERF-FIT was not supported, the possible influence of cultural differences was emphasized. A comparison between Dutch and Brazilian children showed that generalizing gender-specific norms across countries can lead to erroneously diagnosing or failing to diagnose children as being at risk of poor motor skill related fitness. Further research is needed to replicate findings about gender differences in motor skills among Dutch and Brazilian children, and to establish norm scores for other countries in which the PERF-FIT is used.

Keywords: PERF-FIT, gender, motor skills

Gender differences in motor skill performance on the PERF-FIT

Physical activity plays a central role in maintaining good physical health. A lack of it has been shown to be a risk factor of a similar severity as smoking and is assumed to be causally related to coronary heart disease, type 2 diabetes, various types of cancer and premature mortality (Lee et al., 2012). Hence, establishing a lifestyle with adequate levels of physical activity greatly contributes to an individual's overall health status, whereby the amount of physical activity during the childhood years is a crucial factor. Evidence suggests that its benefits extend into adulthood, as active and healthy children as a result are more likely to be physically active later in life (Malina, 2001). A key mechanism behind this carryover effect is the development of motor skills. A child that regularly engages in physical activity gets more opportunities to acquire motor skills through practice. Children with proficient motor skills in turn are more motivated to participate in physical activities and develop into physically active adolescents. In light of this bi-directional relationship, motor skill development during childhood has been proposed as an effective strategy to increase levels of physical activity across the lifespan (Barnett et al., 2009). Thus, it is important to identify children who are less able than their peers to perform motor tasks and therefore at risk of poor motor skill related fitness. To filter out these children, it is necessary to determine which level of motor skills is typical for a child at a certain age, so that PE teachers or other professionals can detect those who perform atypically. However, not only the differences between age groups have to be taken into account but also the differences between males and females, since a number of researchers have suggested that differences in motor skills between genders exist as early as since childhood. As Rodrigues et al. (2019) reported, the typical assumption is that males have better gross motor skills than females, while females have better fine motor skills than males. Gross motor skills are related to large muscles, such as "crawling, walking or running" (Gonzales et al., 2019, p.2), while fine

motor skills like “grasping, object manipulation or drawing” (Gonzales et al., 2019, p.2) rely on small muscle groups. Indeed, Rodrigues et al. (2019) found support for this assumption. They conducted a meta-analysis of 19 studies that used the MABC-2 to examine school-aged children. Boys performed better than girls at gross motor skill tasks in 75% of the studies, while girls showed better fine motor skills than boys in 65% of the studies. Another meta-analysis by Barnett et al. (2016) that reviewed 59 studies also concluded that boys have better gross motor skills than girls.

In line with this, the Bruininks-Oseretsky Test second edition (BOT-2), one of the most widely used motor skill tests for children, uses different norms for each gender. The Movement Assessment Battery for Children second edition (MABC-2) in contrast, does not make a distinction between boys’ and girls’ performance on motor skill tasks. Kokštejn et al. (2018) stated that for three- to four-year-old children, the MABC-2 still offers a valid measurement for both genders. However, starting from the age of five, differences were observed between boys and girls, which is why they suggested gender-specific norms from five years onwards to maintain the validity of the test. Similarly, Rodriguez-Negro et al. (2019) found that the gender differences in motor competence are not observable yet among very young children and emerge only until after the age of eight.

The Performance and Fitness Battery (PERF-FIT)

In 2020, the Performance and Fitness Battery (PERF-FIT) has been developed as a more affordable alternative to enable testing of children in low- and middle- income areas. For the PERF-FIT, preliminary gender-specific norms have been developed based on a sample with children from South Africa, which need to be validated in the next step. The PERF-FIT consists of two subscales: the agility and power subscale and the motor performance subscale. The latter is further divided into the ball skills/object control scale, the balance/stability skills scale and the locomotor skills scale.

Predictions regarding the gender differences on the PERF-FIT subscales

Based on previous research findings it is possible to make predictions on each subscale in regards to which gender will outperform the other. Raudsepp and Pääsuke (1995) found that at the young age of five years, boys already outperformed girls at long throwing and maximal strength tasks. From the age of eight years on, boys also surpassed girls in running speed and lower body explosive strength. Given the high overlap of these measures with the agility and power subscale, which includes measures of running, jumping and throwing, it is thus plausible to assume that boys will perform better than girls on this subscale. Regarding ball skills, it is commonly assumed that boys spend more time with ball games than girls, which has been supported by research on gender differences in leisure activities (Badric et al., 2015). In line with this, Valtr et al. (2016) observed better aiming and catching skills in boys as compared to girls. Since these are key to performing well on the ball skills/object control scale, which measures how well a child can repeatedly bounce, throw and catch a ball, boys will likely outperform girls on this scale as well. Regarding balance skills, a meta-analysis by Rodrigues et al. (2019) revealed that in most studies it was concluded that girls can maintain balance better than boys from an early age. Possible explanations include the fact that girls engage more in activities that encourage the development of balance skills but also the fact that girls are on average shorter than boys, which facilitates balance maintenance (Valtr et al., 2016). Girls are therefore expected to perform better than boys on the balance/stability skills scale, which assesses the ability to maintain balance on one leg while standing or moving. Locomotor skills are measured through tasks such as jumping or hopping and are assumed to be related to static and dynamic balance skills (Smits-Engelsmann et al., 2020). Indeed, Sugden et al. (2013) describe the finding that girls outperform boys in hopping tasks as one of the most distinct differences

between males and females before the age of six. It is thus likely that girls will outperform boys on the locomotor skills scale as well.

This paper focuses on examining the differences between the performances of boys and girls on the PERF-FIT in an attempt to answer the question of whether the use of gender-specific norms is necessary in the case of the PERF-FIT. Additionally, it investigates the effect of age on gender differences on the PERF-FIT to clarify whether differential gender norms are appropriate for all ages.

Hypothesis 1. Scores on the subscales of the PERF-FIT will differ significantly between genders. Males will have higher average scores than girls on the power and agility scale (running, stepping, side jump, long jump and overhead throw) and on the ball skills/object control scale (bouncing and throwing). Females will have higher average scores on the balance/stability skills (static balance and dynamic balance) and on the locomotor skills scale (jumping and hopping (right and left)).

Hypothesis 2. There will be a significant positive interaction between age and gender affecting gender differences in motor skills, meaning that the differences in motor skills between genders will increase with age.

Method

Participants

A total of 249 school-aged children participated in this study, out of which 173 were assessed in the Netherlands; 73 from regular schools and 100 from a school for children with special needs. The other 76 children were assessed in Brazil, all in regular schools. For the recruitment, several schools were contacted and asked for the headmaster's agreement to participation in the study. Those children whose parents had filled in the informed consent form took part in the study. Out of the 249 children, 155 had been diagnosed with a neurodevelopmental or learning disorder. In order to increase the samples' ecological

validity, it was adjusted to resemble the general population in terms of the prevalence of disorders. Therefore, 12 children with ADHD, 22 with ASS, one with OCD, 24 with DCD and 40 with learning disorders were removed. One four-year-old typically developing child was removed from the Dutch sample, since no children in the Brazilian sample were as young. The final sample consisted of 149 boys and girls. Table 1 summarizes the frequencies and percentages of children for each disorder and country. As shown, the number of girls was somewhat larger than the number of boys and the number of Dutch children was substantially larger than the number of Brazilian children. The age range was five to 13 years (M_{age}= 9.2, *SD*= 1.7). This study is approved by the Ethics Committee of Psychology (PSY-1920-S-0107).

Table 1.

Frequencies and Percentage of Neurodevelopmental Disorders and Nationality per Gender

	Neurodevelopmental disorder					Nationality			
	TD	ADHD	DCD	ASS	LD	Dutch	Brazilian	Total	Percent
Boys	40	6	10	1	10	47	20	67	45.0
Girls	53	3	12	1	13	50	32	82	55.0
Total	93	9	22	2	23	97	52	149	100.0
Percent	62.4	6.0	14.8	1.3	15.4	65.1	34.9	100.0	

Note. TD = Typically developing, ADHD = Attention deficit hyperactivity disorder, DCD = Developmental Coordination Disorders, ASS = Autism spectrum disorder, LD = Learning disorder

Materials

Demographic questionnaire

A demographic questionnaire was used to assess the relevant demographic variables. Height and weight were measured. Other variables were reported by the child, such as date of birth, whether a class had been repeated, whether swim diplomas had been obtained, weekly

hours of PE at school, weekly hours of physical activity as well as television and gaming during leisure, types of sport or games during leisure and the presence of disorders or regular use of medication.

EuroQol-5-Dimension-Youth (EQ-5D-Y)

The EQ-5D-Y is a child-friendly version of the EuroQol-5-Dimension (EQ-5D), a questionnaire that assesses quality of life (QoL). It asks the child whether they suffer from any impairments in their ability to walk about, take care of themselves, participate in everyday activities and hobbies and how often they experience pain and negative emotions. The children respond by choosing one of three levels: no problems, some problems or a lot of problems. Finally, each child indicates their current health status on a scale from 0 to 100. The EQ-5D-Y has been found an appropriate tool for the assessment of QoL in children (Wille et al. (2010).

Performance and Fitness Battery (PERF-FIT)

The (PERF-FIT) was used to measure the motor skills of the children. It consists of the agility and power subscale and the motor skills performance subscale. Its use for the assessment of motor skills in children has been supported by several studies that show that the PERF-FIT has adequate construct, content (Smits-Engelsmann et al., 2020) and ecological validity (Doe-Asinyo & Smits-Engelsmann, 2021), as well as inter-rater and test-retest reliability (Smits-Engelsmann et al., 2020) and feasibility in low-resource contexts (Smits-Engelsmann, 2021).

Agility and Power subscale. This subscale contains the items *running, stepping* and *side jump,* which measure agility by assessing the speed with which the child can run or step through the fields of a ladder on the ground. Touching the borders of the fields results in penalty points. The items *long jump* and *overhead throw* require the child to jump or throw as far a possible and measure explosive strength. The participants have at least two trials on

each item. If required, an extra trial for running and stepping are possible. The two children that are assessed together alternate to allow for 15 seconds of rest in between the trials.

Motor skills performance subscale. This subscale is divided into three subscales; ball skills/object control, balance/stability skills and locomotor skills. The ball skills/object control subscale consists of the items *bouncing and catching* as well as *throwing and catching*, which require the child to bounce/throw a ball and catch it again ten times. This is done for five rounds, which become increasingly more difficult by instructing the child to use only one hand or clap in between bouncing/throwing and catching. The balance/stability skills subscale includes the items *static balance*, which measures the time the child can maintain balance standing on one leg, as well as *dynamic balance*, which measures how well the child can maintain balance on one leg while walking or moving cans on the ground. The locomotor skills subscale consists of the *jumping and hopping (left and right) series*. Every child starts at the easiest level and is allowed to proceed to the next level if the previous one is completed successfully after two trials. Only one trial is administered if the maximum score is reached immediately. First, the child is instructed to jump through the fields of a ladder on the ground and land with both legs with out touching the borders. The difficulty increases in the second round, by requiring longer jumps and in the third and fourth round by requiring higher jumps as well. This procedure is repeated with hopping, first on the right, then on the left leg.

Research Design and Procedure

This study was designed as a cross-sectional field study. The participants completed the assessment during school hours in the gym of their school. Two children of similar ages were paired to complete the assessment together. For each pair of children two observers were present to note down the test scores. Each assessment took between 40 and 60 minutes. First, the child answered the questions from the demographic questionnaire and then from the

EQ-5D-Y. Subsequently, the PERF-FIT was completed. Afterwards, the children were rewarded for their effort with a small toy. While most of the children were tested consecutively, a small number of children, who had to interrupt the assessment to attend their PE lesson, completed the test on the following day. Three children had to be re-tested on the locomotor skills subscale because the research assistants had mistakenly instructed them to first complete the items that are most effortful for the leg muscles, which might have affected the subsequent performance on the other items of the subscale.

Statistical Analysis

A Chi-square test was conducted to check for an association between gender and the prevalence of neurodevelopmental disorders. Another Chi-square test was performed to check for an association between gender and the proportions of children from each country. This was done to ensure that any differences between boys and girls were not due to differences between the genders in terms of disorder prevalence or nationality. Assumption for the Chi-square tests were checked, no violations were found. An ANOVA was conducted to check whether the mean age was similar between boys and girls as well as Dutch and Brazilian children, to ensure that the results were not affected by differences in the mean ages. The assumptions for an ANOVA were checked. Since Levene's test ($F(1, 147) = 19.366$, $p = 0.000$) revealed that the two countries had unequal variances in their distributions of age, Welch's test was used.

Four MANOVA's were performed to compare the genders on motor skills, based on the raw scores on each single item of the PERF-FIT (best scores). Items were combined based on the subscales and the construct they measure; *Running* and *stepping* as agility measures, *long jump* and *overhead throw* as power measures, *bouncing* and *throwing* as ball skill measures and *jumping, hopping (right)* and *hopping (left)* as locomotor skill measures. *Dynamic balance* was treated as a single item and therefore analysed with an ANOVA

because *static balance* had to be excluded from all analyses, as it was not suitable for comparison due to different maximum scores in the two countries. The item *side jump* was treated as a single item, since Levene's test revealed a violation of the homogeneity of variances assumption for its test score distribution ($F(1, 147) = 7.062$, $p = 0.009$), so Welch's test was used. Eta-squared (η^2) was chosen as the measure of effect size. To examine the effect of age on gender differences, a regression including the factors gender, age and the interaction between gender and age was conducted for any test item that showed a significant difference between boys and girls. The assumptions for a regression analysis were checked, there were no violations. For all analyses a critical p-value of 0.05 was used.

Results

There was no association in the sample between gender and the likelihood of having a neurodevelopmental disorder, as indicated by the non-significant result of the first Chi-squared test ($\chi^2 = 1.900$, $p = .754$). Also, the proportions of children from the Netherlands and Brazil did not differ by gender ($\chi^2 = 1.366$, $p = .243$). As shown in Table 2, the mean age of the two genders and the two countries was approximately equal (respectively, $F = 0.019$, $p = 0.890$; $U = 2395$, $p = 0.607$).

Table 2.

ANOVA / Welch's test results for the mean age comparisons between the genders and the countries

	Boys	Girls	F	Sig.	Dutch	Brazilian	t	Sig.
M_{age}	9.25	9.29	0.019	0.890	9.28	9.27	2395	0.607
SD	1.72	1.70			1.94	1.15		

Note. M_{age} = Mean age in years, SD = Standard deviation, Sig. = Significance

Comparison between boys and girls based on their raw scores on the PERF-FIT

As can be seen in Table 3, the differences were non-significant for *running* and *stepping, side jump, long jump* and *overhead throw, bouncing* and *throwing, jumping, hopping (right)* and *hopping left*. On *dynamic balance*, the average score was 24.0 for girls and 21.6 for boys, scores that were obtained by combining of the number of steps successfully taken and the number of cans successfully shifted while maintaining balance on one leg. This difference was significant with $F = 4.154$, $p = .043$ with a small effect size of $\eta^2 = .028$.

Table 3.

Results of the gender comparisons with MANOVA / ANOVA / Welch's test for the best scores on the PERF-FIT items

	F	Sig.	η^2
1. Running & Stepping	0.749	.475	.010
2. Side Jump	0.813	.369	.005
3. Long Jump & Overhead Throw	0.346	.557	.002
4. Bouncing & Throwing	0.198	.820	.003
5. Jumping & Hopping (Right, Left)	0.947	.420	.019
6. Dynamic Balance	4.154	**.043**	.028

Note. MANOVA was used for comparison 1, 3, 4 & 5, Welch's test for comparison 2 and ANOVA for comparison 6

The regression analysis that was conducted for the item *dynamic balance* was non-significant ($F(3, 91) = 2.567$, $p = 0.059$) with $R^2 = .078$, which means that none of the factors gender, age or the interaction between gender and age were significant predictors for the raw score on *dynamic balance*.

Post hoc exploration

To explore possible cultural differences in the children's performance on the PERF-FIT, the gender differences were inspected separately within the Dutch and the Brazilian

sample. They were calculated by subtracting the boys' raw scores from the girls' raw scores. Exceptions were made for item 1 (*running*) and 2 (*stepping*), which are the only items on which a lower number indicates a better performance, since they measure speed in seconds. Here, the raw scores of the girls were subtracted from that of the boys. Therefore, on all items, a positive number indicates that boys outperformed girls, while a negative number indicates that girls outperformed boys.

Table 4.

Mean differences between boys and girls on the PERF-FIT single test items per country

	Gender difference Dutch Sample	Gender difference Brazilian Sample
1. Running	0.03 s	0.58 s
2. Stepping	0.42 s	0.99 s
3. Side Jump	-2.5 #	2.0 #
4. Long Jump	1.2 cm	7.5 cm
5. Overhead Throw	0.64 cm	17.3 cm
6. Bouncing	0.2 #	2.9 #
7. Throwing	1.0 #	4.3 #
8. Jumping	-0.6 #	1.1 #
9a. Hopping Right	-2.7 #	2.1 #
9b. Hopping Left	-2.9 #	1.8 #
10b. Dynamic Balance	-3.5 #	0.1 #

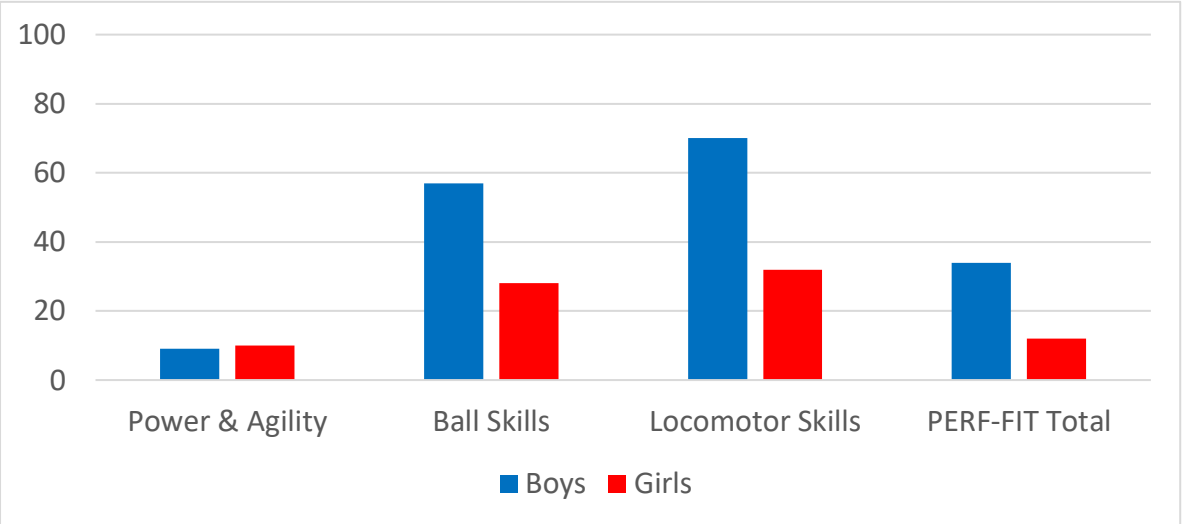
Note. s = seconds, # = number of successful trials, cm = centimetres, positive value = boys outperformed girls, negative value = girls outperformed boys

As shown in Table 4, in the Dutch sample boys scored slightly higher than girls on *running*, *stepping*, *long jump*, *overhead throw*, *bouncing* and *throwing,* while girls scored slightly higher than boys on *side jump, jumping, hopping right, hopping left* and *dynamic balance.* In

the Brazilian sample, boys scored slightly higher than girls on all items. These gender differences were not tested for significance in order to not increase the number of tests performed in this study further. Instead, the percentages of boys and girls classified as being at risk of poor motor skill related fitness were calculated per country, in order to illustrate how the cultural differences would manifest themselves when applying the South African norms to the Dutch and the Brazilian samples. Being at risk was defined as having a converted quartile score of three or lower on the subscales and the overall PERF-FIT. These converted quartile scores were calculated by using the preliminary gender-specific norms from South Africa. The balance/stability skills subscale was excluded, since it contained item 10a (*static balance*), which had been removed from all analyses. As shown in Figure 1 and 2, the percentages of children classified as at risk vary sharply between the genders in the Dutch sample, but less in the Brazilian sample.

Figure 1

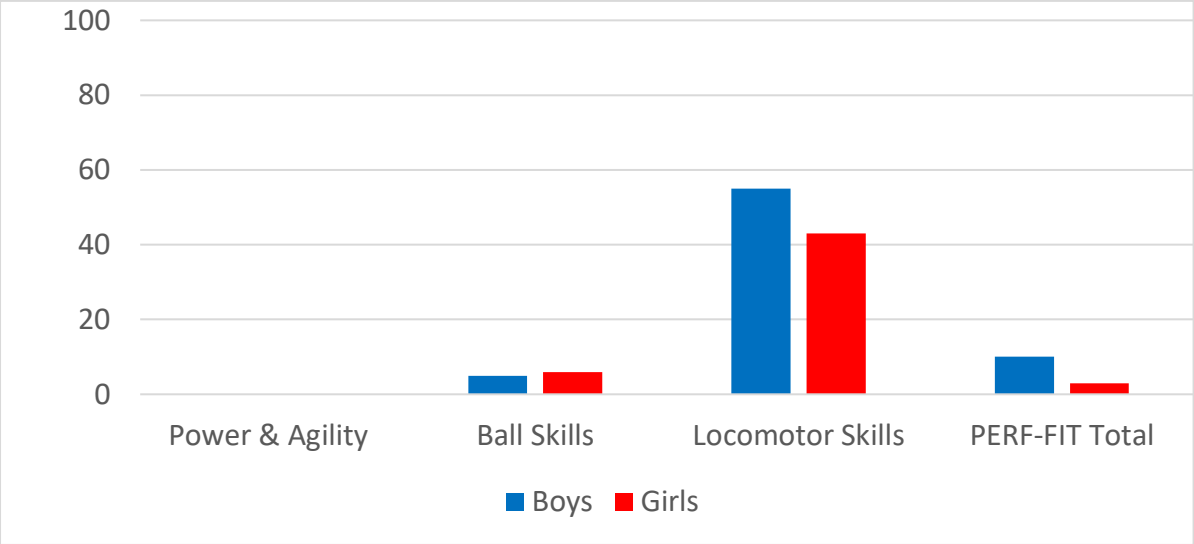
Percentage of Children classified as being at Risk in the Dutch Sample



Note. Being at risk equals having a converted quartile score of 3 or lower. These scores were calculated based on the gender-specific norms developed in South Africa.

Figure 2

Percentage of Children classified as being at Risk in the Brazilian Sample



Note. Being at risk equals having a converted quartile score of 3 or lower. These scores were calculated based on the gender-specific norms developed in South Africa.

Discussion

The aim of this study was to investigate motor skill differences between boys and girls, assessed with the PERF-FIT. This was done to clarify whether gender-specific norms are appropriate in the case of the PERF-FIT. The first hypothesis was that boys would outperform girls on the agility and power as well as ball skills/object control subscales, while girls would outperform boys on the balance/ stability skills as well as the locomotor skills subscales. The results only gave partial support for this hypothesis, with girls performing significantly better than boys on the item *dynamic balance*, which makes up half of the balance skills scale. However, in all other domains there were no significant differences between the genders. Additionally, all effect sizes were below 0.1, which is considered the cut-off point for a small effect size (Miles, J., & Shevlin, M., 2001). The second hypothesis was that there would be a significant positive interaction between the variables gender and age, resulting in increasingly large gender differences as children become older. Since only the comparison on *dynamic balance* yielded a significant difference, the regression analysis was limited to this item. The interaction effect of gender and age was non-significant, hence, the hypothesis was not supported.

An interesting finding of the post hoc exploration regarding cultural variation was that different patterns were observed regarding the gender differences on the test items. In the Dutch sample, boys scored higher than girls on six out of eleven items, while girls scored higher than boys on the remaining five items. In the Brazilian sample however, boys scored higher on all items. It is important to note that these differences were relatively small on average and not tested for significance. Still, the observed trend might have relevant consequences, as illustrated by applying the preliminary gender-specific norms to the present samples. Dutch boys were more likely to be classified as being at risk of poor motor skill related fitness than Dutch girls on the subscale as well as on the overall PERF-FIT. This is explained by the fact that classifications were based on the converted quartile scores, which were calculated by using the preliminary gender-specific norms. These norms were developed with a South African sample, in which boys obtained higher average scores than girls. Therefore, in the case of a boy and a girl with equal raw scores, the converted quartile score of the boy will be lower than that of the girl. As a consequence, Dutch boys were twice as likely as Dutch girls to be classified as being at risk on the ball skills/object control subscale, even though their raw scores were slightly higher than those of the girls (see Table 4; *bouncing* & *throwing*). On the locomotor skills subscale the gender gap was even larger, since on this scale the Dutch boys had lower raw scores than the Dutch girls (see Table 4; *jumping*, *hopping right* & *hopping left).* In the Brazilian sample the percentages of children identified as being at risk did not vary as much between genders as in the Dutch sample. This observation can be explained by the fact that Brazilian boys performed better than Brazilian girls on all test items and that this pattern of results resembles that of the South African sample. The gender differences in Brazilian sample are therefore in accordance with the preliminary gender-specific norms, in contrast to the pattern of gender differences in the Dutch sample, in which boys only performed better than girls on roughly half of the items.

The fact that both research hypotheses were not supported might seem somewhat inconsistent with previous research findings, since in most studies on motor skills in children, relatively consistent differences between boys and girls were reported. The discrepancy however, between the present study and the existent literature on the topic might be explained by the fact that the PERF-FIT has only been developed recently in 2020 and is the first motor skill test that has been developed with the goal of being feasible for middle- and lower-income areas. As a consequence, virtually all previous research was based on the motor skill tests that were available at the time, such as the MABC-2 or the BOT-2. The only authors that have mentioned gender differences on the PERF-FIT before were Smits-Engelsman et al. (2020). In their Brazilian sample, they found that among typically developing children, boys outperformed girls in *running*, but for all other items and groups of children, their gender did not have any impact on their performance. The present study supports the conclusion that overall, boys and girls were more similar than they were different in terms of their performance on the PERF-FIT.

Strengths and Limitations

One of the biggest strengths of this study was the inclusion of children from two countries, the Netherlands and Brazil, which enabled the investigation of cross-cultural aspects of the topic. The same is true for the inclusion of children who suffered from neurodevelopmental or learning disorders, since they make up a substantial part of the population and were present in this sample with a prevalence that closely resembles the general population of school-aged children. Another strength of this study was the testing of children in the gym of their own schools, where they regularly attend PE classes because thereby, the relevance of the natural environment was acknowledged and any unnatural behaviour that was due to the conditions of the testing situation was reduced.

However, the findings of this study have to be seen in light of some limitations as well. Firstly, the item *static balance* had to be excluded, since for the Dutch children the maximum score on this item had been set at 15 seconds, while for the Brazilian children a maximum score of 10 seconds had been used. This also made it impossible to look at the balance/stability skills subscale as a whole. Another limitation was linked to the significance of the results in this study. During the analysis, one p-value showed significance at the alpha level of 0.05. However, this significance should be interpreted with caution, since multiple comparisons were conducted which might have artificially raised the likelihood of encountering a statistically significant result. Additionally, the exploration of cultural differences was added post hoc and without statistical analyses, which is why its results can only be interpreted as a preliminary exploration of possibly relevant variables. A limitation to the regression analysis was the age of the children in the sample that ranged from five to thirteen years. Kokštejn et al. (2018) observed gender differences at the age of five that were not present in three- to four-year-old children. The regression analysis found no effect of age on the gender difference on *dynamic balance.* However, since this sample did not include children younger than five, the possibility that age did make a difference between the three- and four-year-olds as compared to children aged five years or older cannot be excluded. The distribution of disorders among the children posed another limitation for this study. In the Dutch sample, children were sampled from regular as well as special schools and the prevalence of children with disorders was adjusted to resemble the population. In Brazil, however, it is not as common as in the Netherlands to diagnose children and send them to special schools. Therefore, this study relies on the assumption that the Brazilian sample resembled the general population of Brazilian children in terms of disorder prevalence. Finally, the independent observations that are required for the statistical tests could be questioned, since the children in this study were tested pair-wise. The possibility that one

child's performance and behaviour during the assessment affected the second child and vice-versa can therefore not be excluded.

Implications and future research

Despite these limitations, the present research suggests that the gender differences in motor skills that are commonly reported on tests like the MABC-2 or the BOT-2 might not manifest themselves as strongly when using the PERF-FIT as an assessment tool. Although the results of this study do not support the initial hypotheses, their most important contribution may be that they illustrate the relevance of cross-cultural variables for the investigation of motor skills in children. Gender differences in the PERF-FIT scores might vary in magnitude and even in direction between countries, as shown by comparing the Dutch with the Brazilian children. Therefore, using gender-specific norms that were developed in one country to interpret scores in another country might lead to poor motor skill related fitness getting over diagnosed for one gender, while getting under diagnosed for the other. Future research needs to take these cultural differences into account when establishing norm scores for the PERF-FIT. While in some countries gender-specific norms might be useful, they might not be generalizable to other countries, where it could be appropriate to use different gender-specific norms or no gender-specific norms at all. It could also be considered to investigate differences in motor skills between areas within a country, especially for large countries with a lot of cultural diversity.

In conclusion, the present study has not shown support for the appropriateness of gender-specific norms in the sample of Dutch and Brazilian children combined, although the results still need to be replicated in order to be generalizable. However, this study also acknowledges that the challenge of establishing gender-specific norms should in the future be approached by taking into account cultural variability.

References

Badrić, M., Prskalo, I., & Matijević, M. (2015). Aktivnosti u slobodnom vremenu učenika osnovne škole. *Croatian Journal of Education - Hrvatski Časopis Za Odgoj i Obrazovanje, 17*(2), 299. https://doi.org/10.15516/cje.v17i2.1630

Barnett, L. M., M.P.H, van Beurden, E., Ph.D, Morgan, P. J., Ph.D, Brooks, L. O., Ph.D, & Beard, J. R., Ph.D. (2009). Childhood motor skill proficiency as a predictor of adolescent physical activity. *Journal of Adolescent Health, 44*(3), 252-259. https://doi.org/10.1016/j.jadohealth.2008.07.004

Barnett, L. M., Lai, S. K., Veldman, S. L. C., Hardy, L. L., Cliff, D. P., Morgan, P. J., Zask, A., Lubans, D. R., Shultz, S. P., Ridgers, N. D., Rush, E., Brown, H. L., & Okely, A. D. (2016). Correlates of gross motor competence in children and adolescents: A systematic review and meta-analysis. *Sports Medicine (Auckland), 46*(11), 1663-1688. https://doi.org/10.1007/s40279-016-0495-z

Doe-Asinyo, R. X., & Smits-Engelsman, B. C. M. (2021). Ecological validity of the PERF-FIT: Correlates of active play, motor performance and motor skill-related physical fitness. *Heliyon, 7*(8), e07901. https://doi.org/10.1016/j.heliyon.2021.e07901

Gonzalez, S. L., Alvarez, V., & Nelson, E. L. (2019). *Do gross and fine motor skills differentially contribute to language outcomes? A systematic review*. Frontiers Media SA. https://doi.org/10.3389/fpsyg.2019.02670

Kokštejn, J., Musálek, M., & Tufano, J. J. (2018). Construct validity of the movement assessment battery for children-second edition test in preschool children with respect to age and gender. *Frontiers in Pediatrics, 6*, 12. https://doi.org/10.3389/fped.2018.00012

Lee, I. M., Shiroma, E. J., Lobelo, F., Puska, P., Blair, S. N., Katzmarzyk, P. T., & Lancet Physical Activity Series Working Group. (2012). Effect of physical inactivity on major

non-communicable diseases worldwide: an analysis of burden of disease and life expectancy. *Lancet (London, England), 380*(9838), 219–29. https://doi.org/10.1016/S0140-6736(12)61031-9

Malina, R. M. (2001). *Physical activity and fitness: Pathways from childhood to adulthood*. Wiley. https://doi.org/10.1002/1520-6300(200102/03)13:2<162::aid-ajhb1025>3.0.co;2-t

Miles, J., & Shevlin, M. (2001). *Applying regression & correlation : a guide for students and researchers. SAGE.*

Raudsepp, L., & Pääsuke Mati. (1995). Gender differences in fundamental movement patterns, motor performances, and strength measurements of prepubertal children. *Pediatric Exercise Science, 7*(3), 294–304. https://doi.org/10.1123/pes.7.3.294

Rodrigues, P. C., Ribeiro, M., Sousa, A., Lopes, S., & Barros, R. (2019). Performance on the movement assessment battery for children: A systematic review about gender differences. [desempeño en la batería de evaluación del movimiento para niños: Una revisión sistemática sobre las diferencias de género]. *Revista Internacional De Ciencias Del Deporte, 15*(55), 72-87. https://doi.org/10.5232/ricyde2019.05505

Rodríguez-Negro, J., Huertas-Delgado, F. J., & Yanci, J. (2021). Motor skills differences by gender in early elementary education students. *Early Child Development and Care, 191*(2), 281-291. https://doi.org/10.1080/03004430.2019.1617284

Smits-Engelsman, B. C. M., Bonney, E., Cavalcante Neto, J. L., & Jelsma, D. (2020). Feasibility and content validity of the PERF-FIT test battery to assess movement skills, agility and power among children in low resource settings. *BMC Public Health, 20*(1), 1-1139. https://doi.org/10.1186/s12889-020-09236-w

Smits-Engelsman, B. C. M., Smit, E., Doe-Asinyo, R. X., Lawerteh, S. E., Aertssen, W., Ferguson, G., & Jelsma, D. L. (2021). *Inter-rater reliability and test-retest reliability of the performance and fitness (PERF-FIT) test battery for children: A test for motor skill related fitness*. Springer Science and Business Media LLC. https://doi.org/10.1186/s12887-021-02589-0

Smits-Engelsman, B., Cavalcante Neto, J. L., Draghi, T. T. G., Rohr, L. A., & Jelsma, D. (2020). Construct validity of the PERF-FIT, a test of motor skill-related fitness for children in low resource areas. *Research in Developmental Disabilities, 102*, 103663. https://doi.org/10.1016/j.ridd.2020.103663

Sugden, D., Wade, M. G., & Hart, H. (2013). *Typical and atypical motor development* (Ser. Clinics in developmental medicine). Mac Keith Press. Retrieved October 15, 2021

Valtr, L., Psotta, R., & Abdollahipour, R. (2016). Gender differences in performance of the movement assessment battery for Children - 2nd edition test in adolescents. *Acta Gymnica, 46*(4), 155-161. https://doi.org/10.5507/ag.2016.017

Wille, N., Badia, X., Bonsel, G., Burström, K., Cavrini, G., Devlin, N., Egmar, A. C., Greiner, W., Gusi, N., Herdman, M., Jelsma, J., Kind, P., Scalone, L., & Ravens-Sieberer, U. (2010). Development of the EQ-5D-Y: a child-friendly version of the EQ-5D. *Quality of life research : an international journal of quality of life aspects of treatment, care and rehabilitation, 19*(6), 875–886. https://doi.org/10.1007/s11136-010-9648-y

Appendix

Figure 1. *A pair of participants completing the power and agility subscale: while the girl is completing item 2 (stepping), the boy is waiting for his turn.*

Figure 2. *The boy is completing the throwing series of the ball skills/ object control scale*

Figure 3. *The boy is placing the cans into the next square of the ladder while standing on his right leg, as part of the balance subscale.*