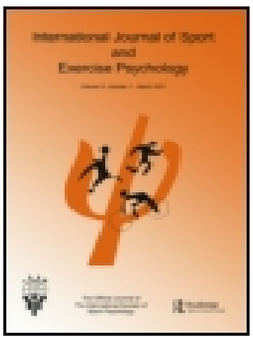

International Journal of Sport and Exercise Psychology

ISSN: 1612-197X (Print) 1557-251X (Online) Journal homepage: http://www.tandfonline.com/loi/rijs20

Reliability and construct validity of the test of gross motor development-2 in Portuguese children

Vítor P. Lopes, Linda Saraiva & Luis P. Rodrigues

To cite this article: Vítor P. Lopes, Linda Saraiva & Luis P. Rodrigues (2016): Reliability and construct validity of the test of gross motor development-2 in Portuguese children, International Journal of Sport and Exercise Psychology, DOI: 10.1080/1612197X.2016.1226923

To link to this article: http://dx.doi.org/10.1080/1612197X.2016.1226923

Published online: 06 Sep 2016.

International Journal of Sport and Exercise Psychology, 2016
http://dx.doi.org/10.1080/1612197X.2016.1226923

Routledge
Taylor & Francis Group

Reliability and construct validity of the test of gross motor development-2 in Portuguese children

Vítor P. Lopes[a,b]*, Linda Saraiva[c] and Luis P. Rodrigues[a,d]

[a]*Research Center in Sports Sciences, Health Sciences and Human Development (CIDESD), Vila Real, Portugal;* [b]*School of Education of Polytechnic Institute of Bragança, Campus Santa Apolónia, Bragança 5300-253, Portugal;* [c]*Escola Superior de Educação do Instituto Politécnico de Viana do Castelo, Viana do Castelo, Portugal;* [d]*Escola Superior Desporto e Lazer de Melgaço, Instituto Politécnico de Viana do Castelo, Viana do Castelo, Portugal*

(*Received 2 March 2016; accepted 11 August 2016*)

The use of several different assessment motor tools make difficult to make comparisons of childhood motor competence across countries. Although the Test Gross Motor Development 2 (TGMD-2) is one of the most used instruments for assessing motor competence, its validation for Portuguese population is needed. The aim of the article is to examine the psychometric proprieties of the TGMD-2, using a Portuguese sample. Totally 330 children aged 5–10 years were assessed with TGMD-2. Cronbach's alpha assessed internal consistency. Test–retest reliability was estimated with the Bland–Altman method. Construct validity was assessed by a confirmatory factor analysis (CFA). The hypothesised model used 12 items and 2 factors: object control and locomotor skills. The test–retest reliability analysis was good, with an agreement ratio of .96 (.09) for 12 skills. Cronbach's alpha values showed acceptable internal consistency (.69 for 12 items, .46 for locomotor skills, and .64 for object control skills). The results of the CFA [CFI = .956, NFI = .868, NNFI = .937, SRMR = .048, and RMSEA = .036 (90% CI: .010–.054)] support the two-factor structure of the original version. Portuguese TGMD-2 version is a reliable and valid tool to assess the gross motor skills of Portuguese children aged 5–10 years.

Keywords: assessment; gross motor skills; validation study; motor development

Introduction

Motor competence is associated with higher levels of physical activity (Holfelder & Schott, 2014; Lopes, Maia, Rodrigues, & Malina, 2011), sport participation (Fransen et al., 2014), physical fitness (Cattuzzo et al., 2016; Haga, 2008), and healthy weight status and body fat percentage (Lopes, Maia, Rodrigues, & Malina, 2012; Lopes, Stodden, Bianchi, Maia, & Rodrigues, 2012; Lubans, Morgan, Cliff, Barnett, & Okely, 2010; Rodrigues, Stodden, & Lopes, 2016). Given the observed global trend to a sedentary lifestyle (Kohl et al., 2012; LeBlanc et al., 2015), with negative consequences for health, it is important to understand the changes associated with motor development and what are the implications of these changes on children's health.

The use of several different motor assessment tools for motor competence makes the comparison challenging of childhood motor performance across countries (Cools, Martelaer, Samaey, & Andries, 2009), impairing a more global understanding of motor competence development. This is partially due to different conceptualisations of motor competence, but also because the assessment instruments used could be process oriented (concerned with the movement pattern) or

*Corresponding author: Email: vplopes@ipb.pt

product oriented (concerned with movement outcome – e.g. distance, accuracy) (Logan, Barnett, Goodway, & Stodden, 2016).

Among several motor assessment tools described in the literature, the Test of Gross Motor Development – second edition (TGMD-2) (Ulrich, 2000) is one of the most widely used instruments in clinical, educational, and research settings. This standardised and norm-referenced tool has been applied to assess the fundamental gross motor skills of children from 3 through 10 years of age, and its normative sample was based on 1208 children residing in 10 states in the United States.

According to Ulrich (2000), the second edition of TGMD provides five principal uses: to identify children who are significantly behind their peers in gross motor skill development; to plan an instructional programme in gross motor skill development; to assess individual progress in gross motor skill development; to evaluate the success of the gross motor programme; and to be used as a measurement instrument in research involving gross motor development.

In fact, over the last decade, the TGMD-2 has been used in several studies to evaluate the motor profile of children with typical development (Hardy, King, Farrell, Macniven, & Howlett, 2010; Pang & Fong, 2009), as well as children with different developmental conditions such as visual impairments (Wagner, Haibach, & Lieberman, 2013), cerebral palsy (Capio, Eguia, & Simons, 2016), autism spectrum disorders (Staples & Reid, 2010), and intellectual disabilities (Rintala & Loovis, 2013). Other studies have applied the TGMD-2 to determine biosocial effects on children's motor development, such as obesity (Cliff et al., 2012), physical activity (Cohen, Morgan, Plotnikoff, Callister, & Lubans, 2014; Foweather et al., 2015), family environment (Barnett, Hinkley, Okely, & Salmon, 2013), perceptions of motor competence (LeGear et al., 2012), and the effectiveness of intervention programmes (Bakhtiari, Shafinia, & Ziaee, 2011; Bardid et al., 2013; Donath, Faude, Hagmann, Roth, & Zahner, 2015; Gursel, 2014; Mostafavi, Ziaee, Akbari, & Haji-Hosseini, 2013).

The test's acceptability among the scientific community is in part due to its easy application and composite structure that allow a multidimensional interpretation of motor development. Empirical evidence of its validity (content, criterion, and construct) and reliability (content sampling, time sampling, and inter scorer differences) is detailed in the TGMD-2 test manual (Ulrich, 2000). Briefly, the manual reports a good internal consistency for each subtest ($\alpha = .76–.92$) and for the gross motor quotient (GMQ) ($\alpha = .87–.94$); acceptable test–retest reliability ($r = .86–.96$ depending on the age group). In regard to the criterion validity, the TGMD-2 has a moderate to strong correlation with the Basic Generalizations subtest of the Comprehensive Scales of Students Abilities ($r = .63$ for locomotor; $r = .41$ for object control, and $r = .63$ for the composite). Finally, Ulrich (2000) reports that the test's construct validity was established by the confirmatory factor analysis (CFA) (chi square/df = 5.29 and Goodness of Fit Index, Adjusted Goodness of Fit Index, and Tucker–Lewis Index ranged from .90 to .96). The validity of TGMD-2 in different groups of individuals (males, females, European Americans, African-Americans, and Hispanic Americans, as well as children with Down syndrome) was also confirmed.

Cross-cultural studies support that TGMD-2 is a valid and reliable instrument to evaluate children's motor performance in Belgium (Simons, Daly, Theodorou, Caron, & Andoniadou, 2008), Brazil (Valentini, 2012), China (Wong & Yin Cheung, 2010), South Korean (Kim, Kim, Valentini, & Clark, 2014), and Philippines (Capio et al., 2016; Capio, Sit, & Abernethy, 2011). To our knowledge, there is no evidence concerning suitability (regional relevance) of TGMD-2 for Portuguese children. Therefore, the purpose of this study was to examine the psychometric properties of the TGMD-2, using a Portuguese sample. The validity and reliability of this tool for the Portuguese population is an important condition for its usefulness in clinical, educational, and research contexts.

Methods

Instrument

The TGMD-2 is a process-oriented instrument, consisting of two subtests with six gross motor skills each: object control skills (catch, striking a stationary ball, stationary dribble, overhand throw, underhand roll, and kick) and locomotor skills (hop, run, gallop, slide, horizontal jump, and leap). It was designed to assess the gross motor development of children 3–10 years old.

The TGMD-2 is a motor assessment tool that requires observational techniques. Each gross motor skill includes three to five behavioural components that are presented as performance criteria. If the child performs a behavioural component correctly, the examiner marks a 1, otherwise marks a 0. Each participant performed one practice trial followed by two trials that were rated. The sum of the observed criteria for each subscale comprises the total raw score (0–48 points). The raw scores can be converted into percentile ranks and standard scores and compared with the ranks of age-matched peers. A GMQ could also be obtained by adding the subtest standard scores and converting the sum to a quotient (Ulrich, 2000).

Instrument translation

The International Test Commission guidelines for translating and adapting tests (AERA, APA, & NCME, 1999; Hambleton, Merenda, & Spielberger, 2005) were followed. All the procedures aimed to guarantee the linguistic, conceptual, operational, and metric equivalence between the Portuguese and original versions. This process comprised two distinct steps: (1) TGMD-2 Translation and Adaptation for the Portuguese language and (2) analysis of the psychometric properties of Portuguese version TGMD-2.

In the first step, the forward-translation design was applied (Hambleton & Patsula, 1999; Peña, 2007). Two authors of this study translated the TGMD-2 Examiner's Manual chapter regarding test structure, administration, and scoring, as well as the Examiner Record Booklet. These authors were fluent in both the languages, familiar with both cultures under study, and knowledgeable in the construction of measuring tools. The comparison of the two translated versions resulted in the first Portuguese TGMD-2 version, which was submitted for evaluation by a panel of Child Motor Development experts (four university professors), for review. The suggestions/corrections proposed by the expert review panel were incorporated into the first Portuguese TGMD-2 version. A pilot study (Saraiva, Santos, Mendes, & Rodrigues, 2007) was conducted to test various aspects of the TGMD-2 administration and scoring, and simultaneously assess the understanding and feasibility for the targeted population. The final Portuguese version displays an identical structure and number of items as the original TGMD-2 version.

In the second step of this study, the Portuguese TGMD-2 version was applied to a sample of 330 children in order to test its reliability (internal consistency and test–retest reliability) and construct validity.

Participants

Five primary schools were approached as a convenience sample and all agreed to participate. Children who met the following criteria were included in the study: (i) Portuguese nationality and (ii) absence of any known intellectual, physical, or emotional disabilities, as well as without special educational needs. Permission was obtained from the respective school director, and parents or guardians gave informed consent and children assented, 330 children (girls $n = 164$); 5–10 years of age (7.9 ± 1.3) participated in the study. Some parents' refusals resulted in

a 98% consent rate. All children were assessed in their attended school. The ethics committee of the Research Centre of Sport, Health and Human Development approved this study.

Data collection procedures

The TGMD-2 was administered according to manual guidelines (Ulrich, 2000) by two researchers, who were specially trained for the task.

The assessment took place in a quiet area of the school with very little intrusions. Depending on the children's age, the individual assessment duration ranged from 30 to 45 minutes. Each skill performance was videotaped for later scoring. One camera was positioned laterally with an angle that permitted the vision of all body movements during the skills execution.

The scoring of each TGMD2 skill was done by two researchers. In order to examine the test–retest reliability, 22 children were assessed twice within 7 days.

Data analysis

Descriptive statistic for TGMD-2 subtests (mean and standard deviation) by age and sex was calculated. Since the normality of the distribution was not guaranteed in every subgroup, the differences between girls and boys in each subtest and by age group were analysed with Mann–Whitney *U* test. The inter-rater reliability analysis using the Kappa statistic was performed to determine consistency among raters, and interpreted as follows (McHugh, 2012): $K < .20$ none; $.21 < K < .39$ minimal; $.40 < K < .59$ weak; $.60 < K < .79$ moderate; $.80 < K < .90$ strong; and $K > .90$ almost perfect.

Test–retest reliability was determined on a subsample ($n = 22$) using the Bland–Altman method (Bland & Altman, 1986). The difference between test and retest measure in each participant was plotted against the mean of the two measures, and the 95% limits of agreement between the two measures were calculated along with the agreement ratio (Nevill & Atkinson, 1997).

Cronbach's alpha was used to analyse internal consistency. The following rules of thumb could be used to interpret the values of Cronbach's alpha (George & Mallery, 2003, p. 231): "$\alpha > 0.9$ – Excellent, $\alpha > 0.8$ – Good, $\alpha > 0.7$ – Acceptable, $\alpha > 0.6$ – Questionable, $\alpha > 0.5$ – Poor, and $\alpha < 0.5$ – Unacceptable", but an alpha of .8 is a reasonable goal.

The construct validity of TGMD-2 was tested using CFA. Robust maximum likelihood (Satorra & Bentler, 2001) was used to estimate model parameters, because the data exhibited a multivariate non-normal distribution (coefficient of Mardia = 15.6). To measure overall fit, we used Standardized Root Mean Square Residual (SRMR), Root Mean Square of Error Approximation (RMSEA), Comparative Fit Index (CFI), Normed Fit Index (NFI), and Non-Normed Fit Index (NNFI). The SRMR represents the average discrepancy between the observed sample and hypothesised correlation matrices, in a well-fitting model the value would be small, equal, or less than .05. The RMSEA measures the degree of misspecification per model degree of freedom, adjusted to the number of estimated parameters in the model (i.e. the complexity of the model). Values below 0.06 indicate an adequacy of model (Hu & Bentler, 1999). The CFI indicates the degree of fit between the hypothesised and null measurement, adjusted to the sample size. The NFI reflects the proportion of the joint amount of data variance and covariance that can be explained by the measurement model being tested. The NNFI is a relative fit index that compares the model being tested to a baseline model (null model), accounting for the degrees of freedom. CFI, NFI, and NNFI values above .95 are considered indicative of a good model fit (Hu & Bentler, 1999).

Statistical programs used were SPSS 21.0 and EQS 6.3.

Results

Table 1 presents the descriptive results of each motor subtest (object control and locomotor) by age group and sex.

A brief data analysis confirmed that the mean values for each motor subtest increased throughout the age group. Independent of sex, older children in both subtests had higher mean scores than younger children. These results support that the TGMD-2 is able to differentiate among the motor development level of children between 5 and 10 years of age. As expected, variability of the results can be seen in both subtests.

In the locomotor subtest, there were no significant differences between boys and girls in any age. On the contrary, in object control the differences occurred in all ages except in 5-year-old children.

Reliability

The Bland–Altman (Bland & Altman, 1986) analysis for test–retest reliability for all TGMD-2 motor skills indicates that the 95% limits of agreement between the 2 measures ranged from 0.80 to 1.13, with an agreement ratio (Nevill & Atkinson, 1997) of 0.96 (0.09). For the locomotor skills, the 95% limits of agreement ranged between 0.85 and 1.17, with an agreement ratio of 1 (0.08). For object control skills the 95% limits of agreement ranged between 0.63 and 1.16 with an agreement ratio of 0.80 (0.13).

The inter-rater reliability analysis varied between 1 for catch, striking a stationary ball, stationary dribble, kick, overhand throw, run, gallop, horizontal jump, leap and slide, and 0.70 for hop and underhand roll.

In terms of internal consistency, Cronbach's alpha value for all 12 items was .69. For the six locomotor skill items, the value was .46. For the six object control skill items, the value was .64.

Construct validity

The TGMD-2 model construct tested for Portuguese sample was similar to that originally proposed by Ulrich (2000). In a two-factor model, locomotion (hop, run, gallop, slide, horizontal jump, and leap) and object control (catch, striking a stationary ball, stationary dribble, overhand throw, underhand roll, and kick) were tested using a CFA.

The obtained model fit indices based on the robust estimation for the Portuguese TGMD-2 model are shown in Figure 1.

Table 1. Mean raw scores and standard deviations ($M \pm SD$) for each TGMD-2 subtest by age group and sex.

		Object control		Locomotor	
Age group	*n*	Girls	Boys	Girls	Boys
5	13 girls; 10 boys	15 (4)	17 (6)	29 (7)	31 (6)
6	18 girls; 16 boys	22 (5)*	27 (6)*	34 (4)	32 (6)
7	39 girls; 49 boys	25 (7)*	30 (7)*	37 (4)	37 (5)
8	40 girls; 33 boys	27 (7)*	32 (7)*	40 (4)	40 (4)
9	24 girls; 36 boys	29 (8)*	34 (6)*	40 (4)	38 (4)
10	30 girls; 22 boys	30 (6)*	38 (5)*	40 (5)	39 (4)

*Significant differences ($p < .05$) between boys and girls; maximum possible score for each subtest is 48 points.

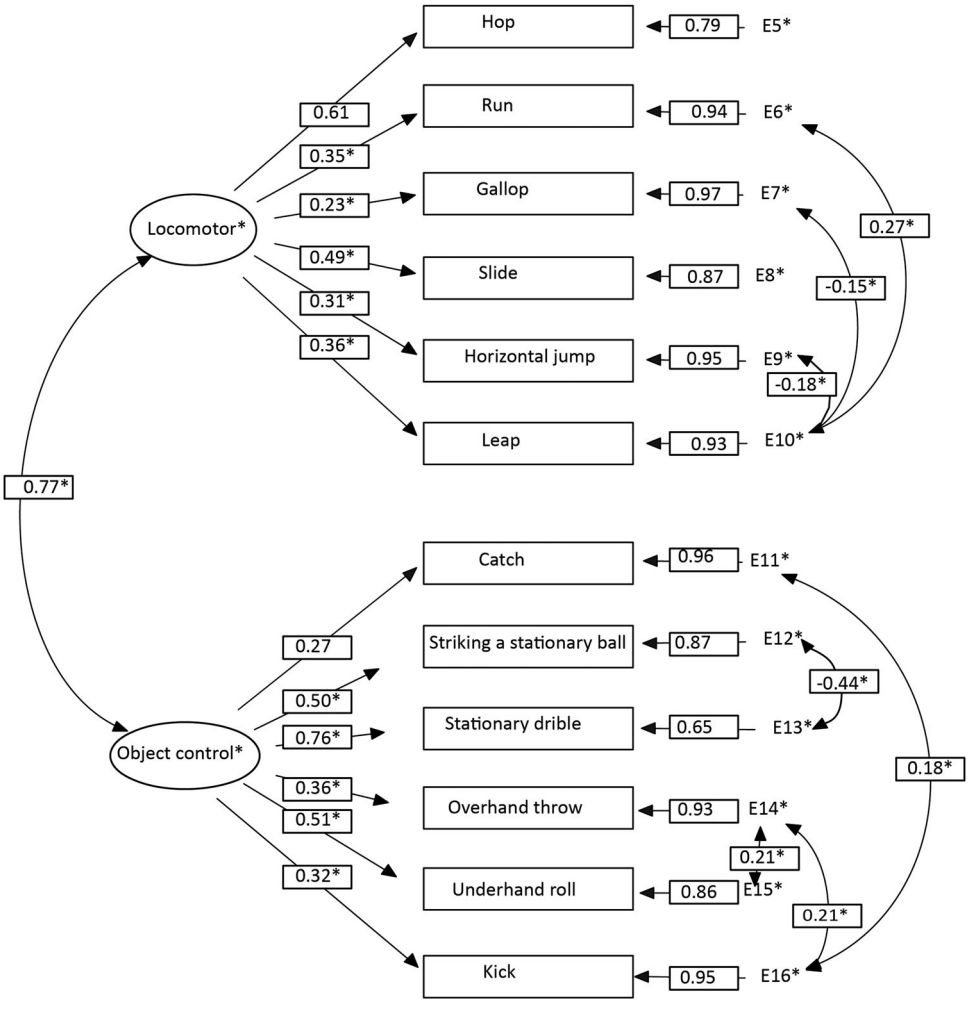

Figure 1. Path diagram of the CFA: two-factor model for Portuguese sample.
Note: the standardised values of item–factor and factor–factor and residual error of each item are presented in diagram.

The values of the goodness-of-fit indices obtained [CFI = .956, NFI = .868, NNFI = .937, SRMR = .048, and RMSEA = .036 (90% CI: .010–.054)] suggest a good fit of the model. In addition, all the loading coefficients were found to be significant (all $p < .05$), with the standardised factor loadings ranging from .31 to .76. A high correlation ($r = .77$; $p < .05$) between the two latent factors was also identified.

Discussion

The development of motor competence in general and fundamental skills in particular is an important aspect linked to children health. Fundamental motor skills form the basis of more advanced and complex motor skills, like the ones used in sports (Gallahue, Ozmun, & Goodway, 2011). Additionally, motor competence is associated with physical activity and health benefits (Robinson

et al., 2015). The assessment of motor competence with a valid and reliable instrument is crucial in motor development intervention programmes. This study contributes to validate such an instrument to Portuguese children. The purpose of the present study was to investigate, in a sample of Portuguese children aged 5–10 years, the reliability and construct validity of the TGMD-2 (Ulrich, 2000), a broadly used instrument to assess gross motor skills competence.

Overall, the results of this study suggest that the Portuguese TGMD-2 version shows similar psychometric characteristics to the original version. Reliability results for internal consistence, inter-rater agreement, and test–retest reliability demonstrate that the TGMD-2 is a reliable instrument for assessing gross motor skills in Portuguese children.

The agreement among different assessors using the same instrument has to be high to guarantee the reliability of the results (Hallgren, 2012; McHugh, 2012). In the present study, the concordance among observers vary from moderate to almost perfect, being in most of items one or close to one, which indicates a very good inter-rater agreement (McHugh, 2012). These results are similar to other validation studies, namely in Brazilian children (Valentini, 2012), South Korean children (Kim et al., 2014), and Flemish children (Simons et al., 2008).

As for test–retest reliability, the values indicate a high temporal stability of the results. Having a reasonable stability over time is important for psychometric instruments (Cook & Beckman, 2006). Other studies (Kim et al., 2014; Valentini, 2012), similar to the original (Ulrich, 2000), also found good values of test–rest reliability.

Concerning to internal consistency, the Portuguese TGMD-2 version revealed an acceptable internal consistency, slightly below the original version (.69 versus .91 for all skills), and studies from South Korea (Kim et al., 2014) and Flemish children (Simons et al., 2008). The lowest value occurred in locomotor skill items reaching an unacceptable value, but since we are doing the validation of the instrument, and for that keeping its initial structure, it is not problematic to have in same cases lower internal consistency values given that the overall fitness-of-good indices remain acceptable. According to Cook and Beckman (2006) reliability, like validity, is a property of the score and not the instrument itself. The same instrument, used in a different setting or with different subjects, can demonstrate wide variation in reliability (Cook & Beckman, 2006; Traub & Rowley, 1991).

As for the construct validity, the results of CFA of the Portuguese TGMD-2 version support the two-factor structure of the original version (Ulrich, 2000).

The Portuguese goodness-of-fit indices were generally good, suggesting that the model fits the data quite well. Indeed, all goodness-of-fit indices were below the criterion value indicative of a good model fit. The two-factor structure is also supported by other validation studies done in Flemish children (Simons et al., 2008), Chinese children (Wong & Yin Cheung, 2010), Brazilian children (Valentini, 2012), South Korean children (Kim et al., 2014), and Filipino children (Capio et al., 2016). Some of the loading coefficients were not high (e.g. Gallop – 0.23 and Catch – 0.27) although found to be significant. This can suggest that some of the skills used were less useful for determining the factor (subscale) for Portuguese children, but since our goal was to provide information on the validation of the intact instrument (in order to be used worldwide), and given the goodness of fit values found, we have to conclude that the inclusion of these tests items did not invalidate the overall fit and validation.

Finally, boys had significantly better performance than girls in object control subtest in all ages, except in 5-year-old children. This tendency was also reported in the literature (Ikeda & Aoyagi, 2008; Saraiva, Rodrigues, Cordovil, & Barreiros, 2013; Thomas & French, 1985), suggesting that boys' advantage in object manipulation skills becomes progressively greater throughout childhood and adolescence. The explanation for this fact has been based in arguments such as social and environmental effects, opportunities for motor experiences, sex stereotyped games and toys, and parental and social expectations (Barnett et al., 2013; Cools, De Martelaer,

Samaey, & Andries, 2011). The TGMD-2 and other standardised tools already consider these sex differences in their norms (e.g. Koöper Koördinationstest für Kinder, Kiphard & Schilling, 1974, 2007; Bruininks-Oseretsky Test of Motor proficiency, Bruininks & Bruininks, 2005), hence, future normative studies with the Portuguese population should also take into account the sexual differentiation in object control subtest.

To our knowledge, this is the first study that sought to explore adaptation and validation of TGMD-2 for the Portuguese population. The different empirical analyses conducted in this study suggest that the Portuguese version shows psychometric characteristics similar to the original version in terms of its sensitivity, reliability, and construct validity, which allows its use in the national context. In the future, it will be important to consolidate this validation process for the Portuguese population. This study should be replicated with a larger sample of Portuguese children from different geographical regions and particularly with ages not explored in this study (from 3 to 5 years).

Conclusion

TGMD-2 is an appropriate tool to assess gross motor skill competence in Portuguese children. The construct of the test battery is supported. Internal consistency for the entire instrument was acceptable. Inter-rater reliability and test–retest reliability were high.

Funding

The authors thank the Portuguese Foundation for Science and Technology for the support grant (UID/DTP/04045/2013 and POCI-01-0145-FEDER-006969).

ORCiD

Vítor P. Lopes http://orcid.org/0000-0003-1599-2180

References

AERA, APA, & NCME. (1999). *Standards for educational and psychological testing*. Washington, DC: Author.

Bakhtiari, S., Shafinia, P., & Ziaee, V. (2011). Effects of selected exercises on elementary school third grade girl students' motor development. *Asian Journal of Sports Medicine*, *2*(1), 51–56.

Bardid, F., Deconinck, F. J., Descamps, S., Verhoeven, L., De Pooter, G., Lenoir, M., & D'Hondt, E. (2013). The effectiveness of a fundamental motor skill intervention in pre-schoolers with motor problems depends on gender but not environmental context. *Research in Developmental Disabilities*, *34*(12), 4571–4581.

Barnett, L., Hinkley, T., Okely, A. D., & Salmon, J. (2013). Child, family and environmental correlates of children's motor skill proficiency. *Journal of Science and Medicine in Sport*, *16*(4), 332–336. doi:10.1016/j.jsams.2012.08.011

Bland, J. M., & Altman, D. G. (1986). Statistical methods for assessing agreement between two methods of clinical measurement. *The Lancet*, *327*, 307–310.

Bruininks, R., & Bruininks, B. (2005). *Bruininks-Oseretsky test of motor proficiency* (2nd ed.). Minneapolis, MN: NCS Pearson.

Capio, C. M., Eguia, K. F., & Simons, J. (2016). Test of gross motor development-2 for Filipino children with intellectual disability: Validity and reliability. *Journal of Sports Sciences*, *34*(1), 10–17. doi:10.1080/02640414.2015.1033643

Capio, C. M., Sit, C. H., & Abernethy, B. (2011). Fundamental movement skills testing in children with cerebral palsy. *Disability and Rehabilitation*, *33*(25–26), 2519–2528.

Cattuzzo, M. T., dos Santos Henrique, R., Ré, A. H. N., de Oliveira, I. S., Melo, B. M., de Sousa Moura, M., … Stodden, D. (2016). Motor competence and health related physical fitness in youth: A systematic review. *Journal of Science and Medicine in Sport*, *19*(2), 123–129. doi:10.1016/j.jsams.2014.12.004

Cliff, D. P., Okely, A. D., Morgan, P. J., Jones, R. A., Steele, J. R., & Baur, L. A. (2012). Proficiency deficiency: Mastery of fundamental movement skills and skill components in overweight and obese children. *Obesity*, *20*(5), 1024–1033. doi:10.1038/oby.2011.241

Cohen, K. E., Morgan, P. J., Plotnikoff, R. C., Callister, R., & Lubans, D. R. (2014). Fundamental movement skills and physical activity among children living in low-income communities: A cross-sectional study. *International Journal of Behavioral Nutrition and Physical Activity*, *11*(1), Article no. 49. doi:10.1186/1479-5868-11-49

Cook, D. A., & Beckman, T. J. (2006). Current concepts in validity and reliability for psychometric instruments: Theory and application. *The American Journal of Medicine*, *119*(2), 166.e167–166.e116. doi:10.1016/j.amjmed.2005.10.036

Cools, W., De Martelaer, K., Samaey, C., & Andries, C. (2011). Fundamental movement skill performance of preschool children in relation to family context. *Journal of Sports Sciences*, *29*(7), 649–660. doi:10.1080/02640414.2010.551540

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.

Donath, L., Faude, O., Hagmann, S., Roth, R., & Zahner, L. (2015). Fundamental movement skills in preschoolers: A randomized controlled trial targeting object control proficiency. *Child: Care, Health and Development*, *41*(6), 1179–1187. doi:10.1111/cch.12232

Foweather, L., Knowles, Z., Ridgers, N. D., O'Dwyer, M. V., Foulkes, J. D., & Stratton, G. (2015). Fundamental movement skills in relation to weekday and weekend physical activity in preschool children. *Journal of Science and Medicine in Sport*, *18*(6), 691–696.

Fransen, J., Deprez, D., Pion, J., Tallir, I. B., D'Hondt, E., Vaeyens, R., … Philippaerts, R. M. (2014). Changes in physical fitness and sports participation among children with different levels of motor competence: A two-year longitudinal study. *Pediatric Exercise Science*, *26*(1), 11–21. doi:10.1123/pes.2013-0005

Gallahue, D., Ozmun, J., & Goodway, J. D. (2011). *Understanding motor development: Infants, children, adolescents, adults*. New York, NY: McGraw-Hill.

George, D., & Mallery, P. (2003). *SPSS for windows step by step: A simple guide and reference, 11.0 update*. Boston, MA: Allyn and Bacon.

Gursel, F. (2014). Inclusive intervention to enhance the fundamental movement skills of children without hearing: A preliminary study. *Perceptual and Motor Skills*, *118*(1), 304–315.

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

Hallgren, K. A. (2012). Computing inter-rater reliability for observational data: An overview and tutorial. *Tutorials in Quantitative Methods for Psychology*, *8*(1), 23–34.

Hambleton, K., & Patsula, L. (1999). Increasing the validity of adapted tests: Myths to be avoided and guidelines for improving test adaptation practices. *Journal of Applied Testing Technology*, *1*, 1–16.

Hambleton, R. K., Merenda, P. F., & Spielberger, C. D. (2005). *Adapting educational and psychological tests for cross-cultural assessment*. Mahwah, NJ: Lawrence Erlbaum.

Hardy, L. L., King, L., Farrell, L., Macniven, R., & Howlett, S. (2010). Fundamental movement skills among Australian preschool children. *Journal of Science and Medicine in Sport*, *13*(5), 503–508.

Holfelder, B., & Schott, N. (2014). Relationship of fundamental movement skills and physical activity in children and adolescents: A systematic review. *Psychology of Sport and Exercise*, *15*(4), 382–391. doi:10.1016/j.psychsport.2014.03.005

Hu, L. T., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal*, *6*(1), 1–55. doi:10.1080/10705519909540118

Ikeda, T., & Aoyagi, O. (2008). Meta-analytic study of gender differences in motor performance and their annual changes among Japanese preschool-aged children. *Japanese Journal of School Health*, *4*, 24–39.

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. doi:10.1080/00222895.2014.914886

Kiphard, E. J., & Schilling, F. (1974). *Korper-koordinations-test fur kinder. KTK. Manual*. Weinheim: Beltz Test GmbH.

Kiphard, E. J., & Schilling, F. (2007). *Körperkoordinationstest fur Kinder* (2nd ed.). Weinheim: Beltz Test GmbH.

Kohl, H. W., Craig, C. L., Lambert, E. V., Inoue, S., Alkandari, J. R., Leetongin, G., & Kahlmeier, S. (2012). The pandemic of physical inactivity: Global action for public health. *The Lancet*, *380*(9838), 294–305. doi:10.1016/S0140-6736(12)60898-8

LeBlanc, A. G., Katzmarzyk, P. T., Barreira, T. V., Broyles, S. T., Chaput, J.-P., Church, T. S., … Group, I. R. (2015). Correlates of total sedentary time and screen time in 9–11 year-old children around the world: The international study of childhood obesity, lifestyle and the environment. *PLoS ONE*, *10*(6), e0129622. doi:10.1371/journal.pone.0129622

LeGear, M., Greyling, L., Sloan, E., Bell, R., Williams, B.-L., Naylor, P.-J., & Temple, V. (2012). A window of opportunity? Motor skills and perceptions of competence of children in Kindergarten. *International Journal of Behavioral Nutrition and Physical Activity*, *9*(1), Article no. 29. doi:10.1186/1479-5868-9-29

Logan, S. W., 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*. Advance online publication. doi:10.1080/02640414.2016.1183803

Lopes, V. P., Maia, J. A. R., Rodrigues, L. P., & Malina, R. M. (2011). Motor coordination as predictor of physical activity in childhood. *Scandinavian Journal of Medicine & Science in Sports*, *21*(5), 663–669. doi:10.1111/j.1600-0838.2009.01027.x

Lopes, V. P., Maia, J. A. R., Rodrigues, L. P., & Malina, R. M. (2012). Motor coordination, physical activity and fitness as predictors of longitudinal change in adiposity during childhood. *European Journal of Sport Science*, *12*(4), 384–391. doi:10.1080/17461391.2011.566368

Lopes, V. P., Stodden, D. F., Bianchi, M. M., Maia, J. A. R., & 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.

McHugh, M. L. (2012). Interrater reliability: The kappa statistic. *Biochemia Medica*, *22*(3), 276–282.

Mostafavi, R., Ziaee, V., Akbari, H., & Haji-Hosseini, S. (2013). The effects of SPARK physical education program on fundamental motor skills in 4–6 year-old children. *Iranian Journal of Pediatrics*, *23*(2), 216–219.

Nevill, A. M., & Atkinson, G. (1997). Assessing agreement between measurements recorded on a ratio scale in sports medicine and sports science. *British Journal of Sports Medicine*, *31*(4), 314–318.

Pang, A. W., & Fong, D. T. (2009). Fundamental motor skill proficiency of Hong Kong children aged 6–9 years. *Research in Sports Medicine*, *17*(3), 125–144.

Peña, E. D. (2007). Lost in translation: Methodological considerations in cross-cultural research. *Child Development*, *78*(4), 1255–1264.

Rintala, P., & Loovis, E. M. (2013). Measuring motor skills in Finnish children with intellectual disabilities. *Perceptual and Motor Skills*, *116*(1), 294–303. doi:10.2466/25.10.pms.116.1.294-303

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

Rodrigues, L. P., Stodden, D. F., & Lopes, V. P. (2016). Developmental pathways of change in fitness and motor competence are related to overweight and obesity status at the end of primary school. *Journal of Science and Medicine in Sport*, *19*(1), 87–92. doi:10.1016/j.jsams.2015.01.002

Saraiva, L., Rodrigues, L. P., Cordovil, R., & Barreiros, J. (2013). Motor profile of Portuguese preschool children on the Peabody Developmental Motor Scales-2: A cross-cultural study. *Research in Developmental Disabilities*, *34*(6), 1966–1973. doi:10.1016/j.ridd.2013.03.010

Saraiva, L., Santos, S., Mendes, R., & Rodrigues, L. (2007). *Validação do 'Test of Gross Motor Development – 2' (TGMD2): um estudo preliminar de validação com crianças portuguesas do 1° ciclo de ensino básico* [Validation of the 'Test of Gross Motor Development - 2' (TGMD2): A preliminary validation study with Portuguese children]. Paper presented at the 3° Seminário Internacional 'Educação Física, Lazer e Saúde – Novas Realidades, Novas Práticas, Braga.

Satorra, A., & Bentler, P. M. (2001). A scaled difference chi-square test statistic for moment structure analysis. *Psychometrika*, *66*(4), 507–514. doi:10.1007/BF02296192

Simons, J., Daly, D., Theodorou, F., Caron, C., & Andoniadou, E. (2008). Validity and reliability of the TGMD-2 in 7–10-year-old Flemish children with intellectual disability. *Adapted Physical Activity Quarterly*, *25*(1), 71–82.

Staples, K. L., & Reid, G. (2010). Fundamental movement skills and autism spectrum disorders. *Journal of Autism and Developmental Disorders*, *40*(2), 209–217.

Thomas, J. R., & French, K. E. (1985). Gender differences across age in motor performance: A meta-analysis. *Psychological Bulletin*, *98*(2), 260–282. doi:10.1037/0033-2909.98.2.260

Traub, R. E., & Rowley, G. L. (1991). An NCME Instructional Module on Understanding Reliability. *Educational Measurement: Issues and Practice*, *10*(1), 37–45. doi:10.1111/j.1745-3992.1991.tb00183.x

Ulrich, D. A. (2000). *Test of gross motor development* (2nd ed.). Austin, TX: PRO-ED.

Valentini, N. C. (2012). Validity and reliability of the TGMD-2 for Brazilian children. *Journal of Motor Behavior*, *44*(4), 275–280. doi:10.1080/00222895.2012.700967

Wagner, M. O., Haibach, P. S., & Lieberman, L. J. (2013). Gross motor skill performance in children with and without visual impairments – research to practice. *Research in Developmental Disabilities*, *34*(10), 3246–3252.

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.