Journal of Motor Behavior
Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/vjmb20

Validity and Reliability of the TGMD-2 for Brazilian Children

Nadia Cristina Valentini [a]

[a] Department of Physical Education, Laboratory of Motor Evaluation and Intervention Research, Federal University of Rio Grande do Sul, Porto Alegre, Brazil

Version of record first published: 02 Aug 2012.

To cite this article: Nadia Cristina Valentini (2012): Validity and Reliability of the TGMD-2 for Brazilian Children, Journal of Motor Behavior, 44:4, 275-280

To link to this article: http://dx.doi.org/10.1080/00222895.2012.700967

Journal of Motor Behavior, Vol. 44, No. 4, 2012

RESEARCH ARTICLE

Validity and Reliability of the TGMD-2 for Brazilian Children

Nadia Cristina Valentini
Department of Physical Education, Laboratory of Motor Evaluation and Intervention Research, Federal University of Rio Grande do Sul, Porto Alegre, Brazil.

ABSTRACT. The Test of Gross Motor Development (TGMD-2) is broadly used in research and clinical settings. The author aimed to translate and investigate the content, criteria, and construct validity and reliability of the TGMD-2 for Brazilian children. The study involved translators, experts, and 3,124 Brazilian children from several states of Brazil. Results confirmed language clarity and pertinence of the TGMD-2. Appropriate indices of the confirmatory factorial validity (root mean square error of approximation = .06; comparative fit index = .88; Tucker-Lewis index = .83; normed fit index = .09; goodness-of-fit index = .98; adjusted goodness-of-fit index = .95), test-retest (values from .83 to .91) and inter- and intrarater reliability were found. Concurrent validity between TGMD-2 and Movement Assessment Battery for Children was weak. The TGMD-2 is a validity and reliability instrument for Brazilian children.

Keywords: assessment, children's motor development, gross motor skills, validation study

Early assessment of children's developmental status provides valuable information to identify motor deficits, and to assist practitioners to properly design programs to improve motor competency (Burton & Miller, 1998; Wiart & Darrah, 2001). However, an appropriate assessment of children's motor development depends on the use of reliable and valid instruments (Bunker, 1989; Burton & Miller; Netelenbos, 2005). It is known that results' generalization of certain assessment scores is limited to the population in which the instrument was validated (Vallerand, 1989). Therefore, the use of norms and standardized scores depend on criterion, content, and construct validity (Wiart & Darrah; Yun & Ulrich, 2002).

The Test of Gross Motor Development (TGMD-2) was originally developed in the United States, but has been translated and validated in different countries (e.g., China, the Netherlands). For example, three studies tested the TGMD-2 validity for typically developing Chinese children (Jing & Hong-Xia, 2007**;** Liang & Li, 2005; Wong & Cheung, 2010**).** Wong and Cheung tested 614 (Typical Develop) TD Chinese children, the results provided reasonable support for the TGMD-2 two factor model (goodness-of-fit index [GFI] = .95; comparative fit index [CFI] = .97). Jing & Hong-Xia (2007) also reported acceptable indices of internal consistency (values from .61 to .92) and test-retest reliability (values from .60 to .87) for the same population. However, it is important to notice that for the Chinese children concerns have been raised about the lower scores in the striking skill and cultural differences between the United States and China scores results and norms (Liang & Li).

Because the TGMD-2 is recognized as a valuable instrument in the identification of motor delays (Bunker, 1989; Burton & Miller, 1998), it has also been used to test children with sensory disability (blind children) and those cognitively impaired. The psychometric proprieties for both populations have been shown; for example, satisfactory reliability results (Cronbach's $\alpha = .82–.86$) and validity evidence (GFI = .88; adjusted goodness-of-fit index [AGFI] = .82) for the TGMD-2 were reported for Flemish children with cognitive disabilities (Simons et al., 2008). In addition, interrater satisfactory reliability (index of correlation coefficient [ICC] = .83; Houwen, Visscher, Hartman, & Lemmink, 2007) and internal consistence (Cronbach's $\alpha = .71–.72$), and acceptable test-retest reliability (ICC = .82–.95; Houwen, Hartman, Jonker, & Visscher, 2010) were reported in children from the Netherlands with visual impairment.

Previously the TGMD-2 validity for a group of children living in Porto Alegre, Brazil (Valentini et al., 2008), was reported. However, considering the sociocultural differences observed across different states of Brazil, estimates that accurately characterize the broad social and cultural diversity within Brazil are yet to be investigated. In the present study I incorporated data from the entire country. Children from 10 different states of Brazil were recruited to better portrait the country multiculturalism. Specifically, I aimed to double-back and reverse translate the TGMD-2 for scientific and clinical equivalence and validate the use of the TGMD-2 in Brazil with respect to content, criteria, and construct validity, as well as for internal consistency and concurrent validity.

Methods

Participants

Four bilingual (2 American English and 2 Brazilian Portuguese native speakers) professional translators participated in the cross-cultural translation process. In addition, three health-related PhD professionals (physical therapist, physical educator, and pediatrician) participated in the process of the Brazilian-Portuguese version of the TGMD-2 (TGMD-2-BR) content validation. The school boards of education from 15 cities from 10 different states in Brazil were contacted and approved the research procedures. Two states were randomly selected from each region of Brazil: North, Northeast, Central-West, South East, and South. A meeting was held with teachers and schools administrators from each city to

Correspondence address: Nadia Cristina Valentini, Department of Physical Education, School of Physical Education, Federal University of Rio Grande do Sul, Rua Felizardo Furtado 750 Bairro Jardim Botânico, Escola de Educação Física UFRGS - Porto Alegre/RS, Brasil. e-mail: nadiacv@esef.ufrgs.br

TABLE 1. TGMD-2-BR Scores, by Sex and Age Group

			Locomotor skills				Object control			
	n		Girls		Boys		Girls		Boys	
Age group (years)	Girls	Boys	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
3	42	52	18.30	6.91	18.60	7.50	13.69	4.04	15.88	4.89
4	62	61	23.47	6.88	23.61	6.53	17.24	4.88	21.90	5.64
5	112	108	26.20	7.16	28.10	6.83	17.78	7.16	24.94	8.17
6	186	173	28.07	6.57	29.09	6.83	20.76	7.49	27.58	7.73
7	190	222	29.51	7.45	31.13	7.76	24.11	7.18	31.97	7.35
8	292	285	29.23	6.69	31.32	6.69	26.75	5.90	34.42	6.28
9	271	266	30.31	6.62	30.88	6.85	28.44	5.90	35.25	6.07
10	167	185	31.16	6.35	31.99	6.74	29.67	6.10	36.82	6.24
Total	1,322	1,352	28.70	7.25	29.91	7.54	24.62	7.68	31.60	8.50

Note. TGMD-2-BR = Brazilian-Portuguese version of the Test of Gross Motor Development.

explain the objectives and relevance of this study. All parents were contacted and consent forms were signed as determined by the institutional review board that approved all the procedures. A total of 2,674 children (1,352 boys and 1,322 girls) from 3 to 10 years old (M age = 7.56 years, SD = 1.91 years) participated in the study. Table 1 depicts detailed information about the TGMD-2-BR scores by sex and age groups.

Instruments

The TGMD-2 is a motor assessment tool that requires observational techniques. It was designed to assess the gross motor development of children from 3 to 10 years old. The test was originally validated in 2000 (Ulrich, 2000). The norms were developed based on data from 1,208 children aged from 3 to 10 years old, living in 10 different states of the United States. The test contains 12 motor skills divided into two subtests: locomotor (run, leap, gallop, hop, jump, and slide) and object control (catch, strike, bounce, over- and underhand throw, and kick). The assessment can typically be completed within 20–30 min. 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 group of children were also assessed using the Movement Assessment Battery for Children (MABC; Henderson & Sugden, 1992). The MABC is designed to assess manual dexterity, ball skills, and balance of children from 4 to 12 years old. Raw score, standardized scores, percentiles, and classification of performance are provided. A sociodemographic questionnaire was also used to collect additional information such as gender, chronological or corrected age, living status, number of family members, and family income. The parents or primary caretakers responded these questions.

The TGMD-2 and the MABC test sections were conducted in the schools, with a week interval between, and took approximately 20 min for each child. All test sections were video recorded for further observation and performance scoring. All parents were informed about the results. For those children identified with low motor scores, information about intervention services was provided.

Procedures

A double-back reverse independent translation procedure was adopted (Vallerand, 1989; Hernandez-Nieto, 2002). This procedure involving four bilingual professional translators required two independent translations from English to Portuguese. After that, two independent translations form Portuguese back to English were completed. The translators did not have access to the original English version of the test. After completing the translation process, all four professionals were invited to a meeting in which all translated versions were compared with the original version of the TGMD-2 (Ulrich, 2000). The two Portuguese versions were revised, and the semantics were adjusted based on unanimous agreement. A final translated and edited scale resulted with the TGMD-2-BR.

A panel of experts involving health-related professionals was intentionally selected to test content validity of the TGMD-2-BR. The experts used a Likert-type scale to independently assess language clarity (responses range from 1 (not clear at all) to 5 (very clear)) and the pertinence (responses range from 1 (not pertinent at all) to 5 (very pertinent)) of all motor items (Neuendorf, 2002). Each expert received a Likert-type scale to score clarity and pertinence of all TGMD-2 motor items. The experts independently scored all the items. The preliminary analyses showed a unanimous concern about one motor criterion item for the kicking motor

skill ("Kicks ball with instep of preferred foot [shoe-laces] or toe"). All experts suggested a small specific change of this specific criterion item. It was unanimously suggested the addition of the phrase "the inside of the foot" as considering the specifics how children in Brazil learn how to kick and play soccer. The motor item was adapted ("Kicks ball with instep of preferred foot [shoe-laces], 'inside' or toe"). The modified criterion was field-tested. A preliminary sample analysis was performed with 450 children (3–10 years old). The kick motor skill was assessed adopting a two-step procedure. The children were assessed following the original protocol of the test. The test was videotaped and analyzed offline. The use of the inside foot when kicking was observed in 65% of the children. In addition, all the children that use the inside foot were retested but this time they were asked to use instep, inside of the foot, and toes. The children (100%) who used inside of the foot were also able to use the toes and instep of the foot. Children were asked to kick again follow their own preference; 98% of those children preferred to use the inside foot. The adaptation suggested by the experts, and field test, was incorporated. The experts used the Likert-type scale to rate the TGMD-2 adapted motor criteria for the investigation of content validity.

A total of 2.674 children were assessed to test the construct validity of the TGMD-2. A group of children ($n = 648$) were retested within a 7–10-day interval by the same evaluator for the test-retest reliability. Concurrent validity was tested using the MABC and TGMD-2 in a group of 161 children (from 4 to 10 years of age). This procedure was adopted to compare the results obtained from the TGMD-2-BR and the MABC. All the tests were videotaped for the offline analyses. The test-retest reliability was conducted to confirm consistency of the TGMD-2 scores items and to provide information about the scale temporal stability (Cicchetti & Rourke, 2004). Three independent experts independently assessed the videotapes for the inter- and intrarater reliability tests.

Data Analyses

Content validity was tested by using the content validity index (CVI) in the scores provided by the panel of experts (Waltz, Strickland, & Lenz, 2010). The test-retest reliability and concurrent validity were analyzed by using Pearson's correlation (Cicchetti & Rourke, 2004). Concurrent validity was also investigated using related t tests. The intra- and interrater scores were analyzed using the ICC (Stephen, 2009). To investigate construct validity of the TGMD-2-BR, confirmatory factor analyses were used. Maximum likelihood on the confirmatory factor analyses using six indices was performed to test the goodness of fit of the two-factor model (Berry, Poortinga, Segall, & Dasen, 1992; Bryman & Cramer, 1999; Hernandez-Nieto, 2002; Hu & Bentler, 1999; Kline & Saggino, 1995; Tucker & Lewis, 1973). To correct the effect of model complexity the root mean square error of approximation (RMSEA) was used (Browne & Cudeck, 1993). The CFI, the Bentler and Bonett (1980) normed fit index (NFI), and the Tucker-Lewis index (TLI) coefficients were used to estimate model discrepancies (Jöreskog & Sörbom, 1993; Marsh, Balla, & McDonald, 1988). The GFI and AGFI were checked to provide information about the proposed model estimate covariance and sample covariance (Holbert & Stephenson, 2002). A correlation under .30 indicates a small correlation, whereas values between .30 and 60 and above .60 indicate, respectively, moderate and large correlations (Hernandez-Nieto, 2002; Waltz, 2010).

TABLE 2. Cross-Cultural Results of TGMD-2 Validation Studies: Factor Loading

	Brazil[a] (N = 2,674)		USA[b] (N = 1,208)		China[c] (N = 626)	
Skill	Loc	OC	Loc	OC	Loc	OC
Run	.46	—	.52	—	.52	—
Gallop	.71	—	.66	—	.66	—
Hop	.66	—	.70	—	.70	—
Leap	.53	—	.49	—	.49	—
Horizontal jump	.53	—	.59	—	.59	—
Slide	.55	—	.69	—	.69	—
Striking stationary ball	—	.69	—	.75	—	.75
Stationary dribble	—	.56	—	.61	—	.61
Catch	—	.59	—	.57	—	.57
Kick	—	.75	—	.65	—	.65
Overhand throw	—	.69	—	.75	—	.75
Underhand throw	—	.45	—	.67	—	.67

Note. TGMD-2 = Test of Gross Motor Development; Loc = locomotor; OC = object control.
[a]Present study. [b]Ulrich (2000). [c]Wong and Cheung (2010).

Results

Content Validity

The concordance results for language clarity of TGMD-2 motor items were all higher than .96, and for pertinence the results were all higher than .89. The CVI for clarity and pertinence among experts was also strong for the test ($\alpha = .93$ for clarity and $\alpha = .91$ for pertinence).

Confirmatory Factor Analyses

Ulrich's (2000) two-factor model (locomotor and object control) was examined in the present study using several measures appropriated to large samples. The results of the RMSEA (.06, 90% confidence interval [.06, .07]), CFI (.88), NFI (.09), TLI (.83), GFI (.98), and AGFI (.95) provided support for the two-factor model. Table 2 shows the results of the loading of the factors for the present study and validation studies from the United States and China.

TABLE 3. TGMD-2-BR Measures of Central Tendency and Statistical Results for Test-Retest Reliability

	Measure of central tendency				Test-retest reliability			
	Pretest		Retest		Correlation		Test *t*	
TGMD-2-BR	*M*	*SD*	*M*	*SD*	*r*	*p*	*t*	*p*
Test	56.49	12.42	56.30	11.95	.90	.001	0.90	.37
Subtests								
Locomotor subtest	29.48	6.13	29.51	5.76	.83	.0001	0.23	.82
Object control subtest	27.00	8.02	26.79	7.82	.91	0001	1.61	.11
Motor skills								
Run	6.24	1.81	6.32	1.76	.80	.001	1.68	.09
Gallop	5.20	1.84	5.25	1.74	.51	.001	0.73	.46
Hop	5.22	1.86	5.16	1.76	.57	.001	0.98	.33
Leap	4.05	1.41	4.03	1.34	.54	.001	0.33	.74
Horizontal jump	3.26	1.83	3.18	1.80	.76	.001	1.49	.14
Slide	5.46	2.54	5.58	2.44	.71	.001	1.65	1.00
Striking stationary ball	5.89	2.28	5.93	2.26	.66	.001	0.50	.61
Stationary dribble	3.99	2.87	4.10	2.78	.90	.001	1.78	.08
Catch	4.23	1.68	4.21	1.64	64	.001	0.42	.68
Kick	4.13	1.99	4.06	1.99	.90	.001	2.00	.06
Overhand throw	3.99	2.36	4.01	2.43	.72	.001	0.28	.78
Underhand throw	4.18	2.20	4.13	2.21	.92	.001	1.55	.12

Note. TGMD-2-BR = Brazilian-Portuguese version of the Test of Gross Motor Development.

Reliability

The Pearson test-retest reliability results showed a strong positive and significant correlation ($r = .90$, $p < .0001$) for the TGMD-2 locomotor test ($r = .83$, $p < .001$) and object control ($r = .91$, $p < .001$) subtests. Significant and positive test-retest correlations were observed for all locomotor skills ($r = .51$–$.92$). Furthermore, nonsignificant differences between the test and the retest scores ($p > .05$) were observed. Table 3 provides the means and standard deviations of the TGMD-2 scores at the test and retest and the *p* values for the statistical analyses.

The interrater reliability for the locomotor ($\alpha = .88$) and object control ($\alpha = .89$) subtests were appropriate. The same trend was observed for locomotor (.86–.94) and object control (.87–.92) skills. The intrarater reliability coefficients ($\alpha = .92$–$.99$) indicated strong and congruent concordance among the experts. Nonsignificant differences among the experts' scores were also observed ($p > .05$).

Concurrent Validity

Scores for the TGMD-2 and MABC were converted to percentile rankings. Person correlations indicated a significant, positive, and small relationship between the TGMD-2 and MABC percentiles ($r = .27$, $p < .001$) for the total sample. The correlation between performance on the TGMD-2 and MABC explained only 7.29% of the variance. Correlations in each age group (each age from 4 to 10 years old) revealed that for the young children, at ages 4 ($r = .42$, $p = .05$) and 5 years old ($r = .56$, $p = .002$), the associations were moderate, explaining 17.6% and 31.4% of the variance, respectively. From ages 6 to 10 years old the correlations were small and nonsignificant ($r = .14$–$.30$, $p > .05$). A related samples *t* test revealed significant differences between both instruments for the total sample, $t(161) = -8.52$, $p < .001$, and all age groups ($ps \leq .007$). Children scored significantly higher on the MABC (*M* percentile = 23.57; *SD* = 24.57) compared with the TGMD-2 (*M* percentile = 7.50; *SD* = 10.23).

Discussion

In this study I conducted a cross-cultural translation of the TGMD-2 and investigated the validity and reliability of the TGMD-2-BR for Brazilian children. The independent translations resulted in a unified and final Portuguese version, the TGMD-2-BR. The double-back and reverse translation, as well as the panel proof to eliminate potential biased translation that usually occurs when an instrument is translated for another culture by only one translator (Hernandez-Nieto, 2002; Vallerand, 1998).

The panel of expert professionals was unanimous in considering the TGMD-2-BR content appropriate to evaluate motor development of children from 3 to 10 years old. The CVI for clarity and pertinence showed a strong coherence among the experts, as confirmed by the kappa concordance coefficient, which demonstrated correspondence on their answers (Bryman & Cramer; Hernandez-Nieto, 2002). These results indicated that the TGMD-2-BR version had great content validity indexes and clear and pertinent motor

criteria values superior to .80 (Cronbach, 1989). The results emphasized the proper translation of all the items (values $> .89$) related to scale concepts and the theoretic relevance (Bryman & Cramer; Cronbach; Hernandez-Nieto).

Construct validity was measured by using confirmatory factor analyses (Berry et al., 1992; Taylor, Bagby, & Parker, 2003). Each motor skill correlated satisfactory with the correspondent subtest (locomotor [run, gallop, leap, hop, jump, and slide] and object control [strike, dribble, catch, kick, throw, and roll a ball]). Negative correlations were not observed between motor skills and subtests, a result that provides addition support for the two-factor model (Cronbach, 1989). We used multiple indexes of fit were used because each parameter encompassed different strength and weakness of the model adjustability (Taylor et al.; Watkins, 1989). All the results were appropriated.

The result for the RMSEA was .06, and values less than .08 were considered as an indication of reasonable error of approximation (Browne & Cudeck, 1993; Hu & Bentler, 1999). The RMSEA result was similar to the results of a previous TGMD-2 study with Chinese children (Wong & Cheung, 2010). The result form the CFI (.88) indicated a good fit of the scores of the TGMD-2-BR into the two-factor model (Bentler & Bonett, 1980; Jöreskog & Sörbom, 1993). The NFI obtained in the present study (.09) and the TLI (.83) confirmed the fit of the two-factor model (Bentler & Bonett; Marsh et al., 1988). The GFI values were superior to .85 and AGFI was superior to .80. This provided additional support for the fit of the two-factor model (Marsh et al.; Tucker & Lewis, 1973).

The results obtained for the fit of the model analyses were similar to the ones reported by Ulrich (2000) with American children (GFI = .96; AGFI = .95; TLI = .90). Cross-cultural validation studies in China (Wong & Cheung, 2010), Belgium (Simons et al., 2008), and south Brazil (Valentini et al., 2008) also reported a similar trend. Considering that a test construct is expressed by the extent to which the underlying items of the test can be identified and how its traits reflect the model (Anastasi & Urbina, 1997), and the overall fit of the data in the present study provided empirical support for the two-factor model for the Brazilian children.

All reliability results were shown to be adequate (Vallerand, 1998). Concordance among experts using the same instrument must be high and positive to guarantee reliability and validity of the results (Hammond, 2006). Values higher than .80 are considered to indicate sufficient concordance (Hammond; Stephen, 2009). The results from the inter- and intrarater analyses from the present study were appropriated (values from .86 to .94) even when a very conservative approach was adopted. The high reliability indexes observed in the Brazilian validation was similar to the results originally reported by Ulrich (2000) for North American children, and by Wong and Cheung (2010) for Chinese children.

With respect to test-retest reliability, positive, strong, and significant correlations were found between the test-retests scores for TGMD-2-BR test as well as for the locomotor and object control subtests. The motor skills test-retest correlations were also positive and significant, and three moderate correlations were found (hop, leap, and jump); strong correlations were found for all other motor skills. The test-retest reliability was also confirmed by nonsignificant results between the test and retest comparisons. The acquiescence phenomenon (positive or negative) was not detected, also indicating that the data were reliable (Waltz et al., 2010). All the results highlighted the temporal stability of the TGMD-2-BR scores, a crucial reliability measure for an instrument (Bryman & Cramer, 1999; Cronbach, 1989; Hernandez-Nieto, 2002; Kline & Saggino, 1995). It is important to emphasize that one basic measure of reliability for psychometric instruments is to have a reasonable level of temporal stability that can be related to the defining measures of the constructs (Cicchetti & Rourke, 2004). Similar results were previously reported by us on a smaller sample of Brazilian children living in the south of Brazil (Valentini et al., 2008).

The results revealed a significant and small association between TGMD-2 and MABC scores for total sample (.27). Moderate and significant correlations between TGMD-2 and MABC scores were found only for 4- and 5 year-old children. Although these results were satisfactory, concurrent validity should be stronger (Cronbach, 1989). Previous research reported low to moderate correlations (values from .13 to .40) between the TGMD-2 and the MABC for preschool children (Logan, Robinson, & Getchell, 2011). The result of the present study may be related to the purpose of each assessment. The MABC was designed to identify motor impairments in everyday tasks whereas the TGMD-2 was designed to identify children who have delays in fundamental motor skills. It is important to note that acceptable concurrent validities have been reported between TGMD-2 and the Children's Activity and Movement in Preschool Study (CHAMPS) Motor Skill Protocol (Williams et al., 2008) and the Fundamental Motor Skill POLYGON test (Zuvela, Bozanic, & Miletic, 2011). Both studies found high coefficients of association (.94 and .82, respectively). Future studies should the relationship between TGMD-2 and tests that measure motor skill quality of performance (e.g., the Preschooler Gross Motor Quality Scale [Sun, Zhu, Shih, Lin, & Wu, 2010]).

Several procedures adopted in the present study emphasize the efficiency of the TGMD-2-BR to screen motor development of Brazilian infants. All the procedures adopted in the present study underline the importance to ensure that the TGMD-2 is suitable for testing children from different cultural background. Appropriate results were found for content and construct validity, as well as for the reliability. Concurrent validity needs further investigation. Clinical assumptions about the use interchangeable of TGMD-2 and MABC should be carefully depicted at ages when the associations, although positive, were small and nonsignificant. The results of the present study may positively impact daily practice of educators and therapists as well as researchers, as they can rely on a validated and reliable instrument to assess children's development and to design an intervention for Brazilian children.

ACKNOWLEDGMENTS

This work was supported by the Brazilian government: CAPES (Coordenação de Aperfeiçoamento de Pessoal Docente) and CNPq (Conselho Nacional de Pesquisa).

REFERENCES

Anastasi, A., & Urbina, S. (1997). *Psychological testing* (7th ed.). Englewood Cliffs, NJ: Prentice-Hall.

Bentler, P. B., & Bonett, D. G. (1980). Significance tests and goodness of fit in the analysis of covariance structures. *Psychological Bulletin*, *88*, 588–606.

Berry, J. W., Poortinga, Y. H., Segall, M. H., & Dasen, P. R. (1992). *Cross-cultural psychology: Research and applications*. Cambridge, England: Cambridge University Press.

Browne, M. W., & Cudeck, R. (1993). Alternative ways of assessing fit. *Sociological Methods and Research*, *21*, 230–258.

Bryman, A., & Cramer, D. (1999). *Quantitative data analysis with SPSS release perceived symptoms: Criterion validity of the Toronto Alexithymia LOS8 for Windows: A guide for social scientists*. New York, NY: Routledge.

Bunker, L. K. (1989). Review of the test of gross motor development. In J. Conoley & J. Kramer (Ed.), *The tenth mental measurements yearbook* (pp. 843–845). Lincoln, NE: University of Nebraska Press.

Burton, A. W., & Miller, D. E. (1998). *Movement skill assessment*. Champaign, IL: Human Kinetics.

Cicchetti, D. V., & Rourke, B. P. (2004). *Methodological and biostatistical foundations of clinical neuropsychology and medical and health disciplines*. London, England: Taylor & Francis Group.

Cronbach, L. J. (1989). Construct Validity after thirty years. In R. L. Linn (Ed.), *Intelligence: Measurement, theory, and public policy* (pp. 147–171). Bloomington, IL: University of Illinois.

Hammond, S. (2006). Using psychometric test. In G. M. Breakwell, S. Hammond, C. Fife-Schaw, & J. A. Smith (Eds.), *Research methods in psychology* (pp. 182–209). London, England: Sage.

Henderson, S. E., & Sugden, D.A. (1992). *Movement Assessment Battery for Children: Manual*. London, England: Psychological Corporation.

Hernandez-Nieto, R. (2002). *Contributions to statistical analysis*. Mérida, Mexico: Los Andes University Press.

Holbert, R. L., & Stephenson, M. T. (2002). Structural equation modeling in the communication sciences, 1995–2000. *Human Communication Research*, *28*, 531–551.

Houwen, S., Hartman, E., Jonker, L., & Visscher, C. (2010). Reliability and validity of the TGMD-2 in primary-school-age children with visual impairments. *Adapted Physical Activity Quarterly*, *27*, 143–159.

Houwen, S., Visscher, C., Hartman, E., & Lemmink, K. A.P.M. (2007). Gross motor skills and sports participation of children with visual impairments. *Research Quarterly for Exercise and Sport*, *78*(2), 16–23.

Hu, L. T., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equations Modeling*, *6*(1), 1–55.

Jing, L .I., & Hong-Xia, M. A. (2007). Study of the credibility and validity of the test of gross motor development of children. *Journal of Physical Education*, *14*(3), 11–31.

Jöreskog, K. G. (1993). Testing structural equation models. In K. A. Bollen & J. S. Long (Eds.), *Testing structural equation models* (pp. 294–316). London, England: Sage.

Kline, P., & Saggino, A. (1995). Item factor analysis of the Italian version of the Myers-Briggs Type Indicator Source. *Journal of Personality and Social Psychology*, *19*, 243–249.

Liang, G., & Li, J. (2005, June). *Construct validity of the test of gross motor development: A cross culture validation approach*. Paper presented at the American Alliance of Health, Physical Education, Recreation and Dance Convention, Chicago, IL.

Logan, S. W., Robinson, L. E., & Getchell, N. (2011). The comparison of performance of preschool children on two motor assessments. *Perceptual and Motor*, *113*, 715–723.

Marsh, H. W., Balla, J. R., & McDonald, R. P. (1988). Goodness-of-fit indexes in confirmatory factor analysis: The effect of sample size. *Psychology Bulletin*, *103*, 391–410.

Netelenbos, J. B. (2005). Teacherŕ ratings of gross motor skills suffer from low concurrent validity. *Human Movement Science*, *24*, 116–137.

Neuendorf, K. A. (2002).*The content analysis guidebook*. London, England: Sage.

Simons, J., Daly, D., Theodorou, F., Caron, C., Simons, J., & 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.

Stephen, J. W. (2009). *Quality of life outcomes in clinical trials and health-care evaluation: A practical guide to analysis and interpretation*. London, England: Antony Rowe.

Sun, S. H., Zhu, Y. C., Shih, C. L., Lin, C. H., & Wu, S. K. (2010). Development and initial validation of the Preschooler Gross Motor Quality Scale. *Research in Developmental Disabilities*, *31*, 1187–1196.

Tucker, L. R., & Lewis, C. (1973). The reliability coefficient for maximum likelihood factor analyses. *Psychometrika*, *38*(1), 1–10.

Taylor, G. J., Bagby, R. M., & Parker, J. D. A. (2003). The 20-Item Toronto Alexithymia Scale IV. Reliability and factorial validity in different languages and cultures. *Journal of Psychology*, *55*, 277–283.

Ulrich, D. (2000). *The test of gross motor development*. Austin, TX: Prod-Ed.

Valentini, N. C., Barbosa, M. L. L., Cini, G. W., Pick, R. K., Spessato, B. C., & Balbinotti, M. A. A. (2008). Test of gross motor development: Expert validity, confirmatory validity and internal consistence. *Brazilian Journal of Kinanthropometry and Human Performance*, *10*, 399–404.

Vallerand, J. R. (1989). Toward a methodology for the transcultural validation of psychological questionnaires: Implications for research in the French language. *Canadian Journal of Psychology*, *30*(4), 23–38.

Waltz, C. F., Strickland, O., & Lenz, E. R. (2010). *Measurement in nursing research*. New York, NY: Springer.

Watkins, D. (1989). The role of confirmatory factor analysis in cross-cultural research. *International Journal of Psychology*, *24*, 685–701.

Wiart, L., & Darrah, J. (2001). Review of four tests of gross motor development. *Developmental Medicine & Child Neurology*, *43*, 279–285.

Williams, H. G., Pfeiffer, K. A., O'Neill, J. R., Dowda, M., McIver, K. L., Brown, W. H., & Patel, R. R. (2008). Motor skill performance and physical activity in preschool children. *Obesity*, *16*, 1421–1426.

Wong, A. K. Y., & Cheung, S. Y. (2010). Confirmatory factor analysis of the Test of Gross Motor Development-2. *Measurement in Physical Education and Exercise Science*, *14*, 202–209.

Yun, J., & Ulrich, D. A. (2002). Estimating measurement validity: A tutorial. *Adapted Physical Activity Quarterly*, *19*(1), 32–47.

Zuvela, F., Bozanic, A., & Miletic, D. (2011). POLYGON—A new fundamental movement skills test for 8-year-old children: Construction and validation. *Journal of Sports Science and Medicine*, *10*(1), 157–163.

Received September 17, 2011
Revised May 2, 2012
Accepted June 5, 2012