

Motor assessment instruments and psychometric procedures: A systematic review

Pâmella de Medeiros^{1*}, Marcela Almeida Zequinão², Fernanda Cerveira Fronza¹, João Otacilio Libardoni dos Santos³, Fernando Luiz Cardoso¹

REVIEW ARTICLE

ABSTRACT

It was our objective to identify the psychometric elements to an epistemological reflection through a systematic review of cross-cultural validation procedures of TGMD-2 batteries, MABC-2 and KTK. Searches were carried out by two evaluators independently without year and language restrictions in six databases: Web of Science, Science Direct, Lilacs, Scopus, Pubmed and The Scientific Electronic Library Online - SciELO. The key words used were: "MABC", "TGMD" and "KTK" all of them combined with the word "validity". There was a total of 734 articles, of which, after the exclusion criteria, remained only 11 studies. It was found that there are differences between the authors in relation to the psychometric factors taken into account in cross-cultural validation. So that there was a lack of unanimity of the validation criteria of all studies in this field.

Keywords: Epistemology, Psychometrics, Motor Assessment.

INTRODUCTION

Motor assessment is an excellent technique for the human development professionals to identify the aspects regarding which the individuals assessed need further attention (Payne & Isaacs, 2007). This process is considered as a means to verify the extent to which the objectives are being reached, identifying children who need individual attention and reformulating the work through the adoption of procedures that permit remedying the shortcomings identified in the motor learning process. Hence, this knowledge on the motor performance provides relevant information to plan Physical Education and Sports programs, as it permits the formulation of motor intervention strategies in the programming of daily routines during pedagogical practice (Krebs, Duarte, Nobre, Nazario, & Libardoni dos Santos, 2011).

Due to the importance of motor assessment, for diagnostic purposes as well as to determine intervention strategies, over the years, measures

have been constructed/validated in different countries around the world and have been used in research to assess the motor performance of children and adolescents (Henderson, Sugden, & Barnett, 2007; Piek, Gasson, & Barrett, 2002; Rosa Neto, 2002; Souza, Ferreira, Catuzo, & Correa, 2007; Ulrich, 2000). Among these motor batteries, the "Test of Gross Motor Development Second Edition -TGMD-2" (Ulrich, 2000), the "Movement Assessment Battery for Children - MABC2" (Barnett, Henderson, & Sugden, 2007) and the "Körperkoordinationstest Für Kinder - KTK" (Kiphard & Schilling, 1974) can be highlighted. Although different motor batteries exist, ranging from the eldest to the most recent examples, the above batteries were chosen in function of the large number of studies produced that have used these tests in recent years. In addition, these instruments tend to be the most used in research on motor performance in different countries (Freitas et al., 2016; Liu,

Manuscript received at January 9th 2015; Accepted at September 9th 2016

¹ University of the State of Santa Catarina, UDESC, Florianopolis, Brazil

² University of Minho do Minho, UM, Braga, Portugal

³ Universidade Federal do Amazonas - UFAM

* Corresponding Author: University of the State of Santa Catarina – UDESC. Street Simone Paschoal, nº358. Cep:88080-350 - Coqueiros, Florianopolis – SC, Brasil.. E-mail: pamellademedeiros@hotmail.com

Hamilton, & Smith, 2015; Martins, Silva, Marinho, & Costa, 2015; Rudd et al., 2015), including Brazil (Luz et al., 2015; Moreira, Magalhães, Dourado, Lemos, & Alves, 2014; Van Keulen, Benda, Ugrinowitsch, Valentini, & Krebs, 2016).

Although these tools are widely used to assess the motor performance around the world, few of them have been validated or adapted to the reality they are applied in. In addition, it has been observed that validations or cultural adaptations have not been executed with due psychometric rigor. In that sense, the International Test Commission (ITC, 2010) highlights that many of the measures used in other countries do not apply all recommended procedures for validation and adaptation in other populations. As a result, the batteries diverge on the psychometric elements used (validity and reliability) (Pasquali, 2010), making the instrument weaker (Libardoni dos Santos, 2014). That may be so as many researchers still mix up the conceptual differences between validating and adapting a measure.

When distinguishing between validation and cross-cultural adaptation, it is verified that the cross-cultural adaptation simply translates the original or, exceptionally, compares the translation literally with a back-translation (Reichenheim & Moraes, 2007). The cross-cultural validation of a tool, on the other hand, represents the extent to which the instrument can assess what it intends to and the extent to which this specific test can measure the same thing in different people who do not belong to the population it was created for. For this procedure to be valid and reliable, the psychometric procedures suggested in the literature (Pasquali, 2010).

For the cross-cultural validation of an instrument, some validity and reliability aspects should be taken into account: 1) content validity – intended to define the items to be included in the instrument, with a consistent theoretical framework; 2) construct validity – described based on theoretical concepts and should be transformed into measurable items; and 3) criterion validity – refers to the correlation between the scores of one test and other

performance measures (Libardoni dos Santos, 2014; Pasquali, 2010). Concerning the reliability criteria: 1) test-retest – the instrument is applied twice to the same group of people with an interval between the applications, verifying the reliability between the results of the two applications (Sampieri, Collado, & Lucio, 1996), 2) inter or intra-rater – reproducibility test using more than one rater (inter) and a single rater (intra) to obtain the repeatability of the method (Ribeiro, Trombini-Souza, Iunes, & Monte-Rasso, 2006) and 3) internal consistency – refers to the extent to which the items of a scale are mutually correlated (Sun et al., 2007).

The lack of convergence in the criteria used to validate the instruments can alter their essence as, in the creation as well as in the validation, the instrument should seek cultural clarity, pertinence, reliability and validity (Cicchetti & Rourke, 2004). A non-reliable assessment inevitably leads to mistaken interpretations of individual development (Yun & Ulrich, 2002), resulting in the production of false indicators or failing to detect some kind of disorder and/or motor impairment (Henderson et al., 2007). Some studies present different motor performance results when considering the same population sample, that is, individuals from the same age and the same regions (Berleze, Haefner, & Valentini, 2007; Souza et al., 2007; Souza, Berleze, & Valentini, 2008). We speculate that these divergences can be due to the use of instruments with constructs and dimensions that differ from the conditions they are applied in, often not demonstrating validity for the reality of the population it is applied in.

The weakness of the motor batteries concerning the validations and adaptations to other cultures, in combination with their extensive use all over the world, indicate a great problem in the field of motor performance and in Physical Education. These aspects make it difficult to compare results among the studies, and often end up underestimating or overestimating the performance of children and adolescents in countries in which no adequate validation was undertaken. That reality points towards possible interferences and, consequently, discrepant results, in research as well as in education,

considering that researchers are often unable to establish a point of normality for the study population, and that interventions by Physical Education teachers can be impaired by assessment that do not express the students' reality.

Thus, as the authors are unfamiliar with other studies intended to join further details on this theme, this review emerged from the need to verify if the studies are following the appropriate psychometric rigor for the validation, with a view to improving the quality of new studies, and with a view to enhancing a discussion among motor performance researchers about possible solutions to the gaps in the literature. Then, we can demonstrate what psychometric elements should be followed, granting support for further validations and even for the construction of new instruments. Thus, the objective in this study was to confront information from the literature about the psychometric elements taken into account in the cross-cultural validation of the TGMD-2, MABC-2 and KTK batteries in different countries, by means of a systematic review; as well as to elaborate an epistemological reflection on the criteria used to elaborate and validate these instruments.

The motor batteries analyzed: TGMD-2, MABC-2, KTK

The objective of the TGMD-2 (Ulrich, 2000) is to assess the motor behavior, broad functioning and motor development level of children between 3 years and 10 years and 11 months of age. The test consists of multiple fundamental motor skills and assesses how the children coordinate the trunk and limbs while performing a motor skill, that is, it analyzes the presence or absence of different skill components instead of mainly assessing the end product of the performance. The test assesses 12 basic motor skills, divided in two specific components: locomotion skills (run, gallop, hop, leap, horizontal jump and slide); and object control skills (striking a stationary ball, stationary dribble, catch, kick, overhand throw and underhand roll). Both components have 24 performance criteria each and, if the child presents the efficiency criterion correctly, (s)he receives score "1"; if not, (s)he receives "0" in

each of the attempts. The child is scored according to the age and sex for each subtest, determining his/her development level, qualified as a gross motor quotient, which is classified in seven categories: "very poor, poor, below average, average, above average, superior and highly superior".

The MABC 2 (Henderson et al., 2007) is used to identify motor difficulties in children and adolescents between 3 and 16 years of age, consisting of three sets of tasks appropriate to the specific age bands: Age band 1 (3 to 6 years); Age band 2 (7 to 10 years), used for analysis in this study, and Age band 3 (11 to 16 years). The tasks to assess each skill in the motor test for Age Band 2 (7 to 10 years) include: Manual Dexterity (placing pegs, threading beads and bicycle trail), Aiming and Catching (aiming and catching with both hands and catching beanbag), Balance (balancing on the table, heel to toe walking and jumping on mats). After the application of the tests, the gross scores are converted into standard scores. The standard scores are added up in each skills category, resulting in the total score of the motor components. By adding up the latter, the standard test score or total result is obtained. Both the standard scores and the total results are compared to a table of percentiles, which permits ranking the children's motor performance. A higher score indicates a higher standard score and, consequently, a higher percentile. Thus, the component score and the total test score and its respective standard scores and percentiles are directly proportional.

The KTK (Kiphard & Schilling, 1974) is a homogeneous battery to assess the bodily coordination capacity intended to examine the basic motor function of children and adolescents between 6 and 14 years of age. Its components are balance, rhythm, laterality, speed and agility, distributed in four tasks, all of which are intended to categorize facets of total body coordination and body mastery. The tasks include: Keeping balance walking backwards; One-legged hopping; Jumping from side to side; and Moving sideways. The KTK uses the same coordination tasks for different ages. Therefore, the contents of the tasks should add difficulties according to the individuals' age. The distinction by age, for

example, is achieved according to criteria like: increased height or distance; increased speed; greater precision in the execution, measured, for example, in function of the larger number of hits in a certain number of attempts. The test also permits different types of presentations and discussion of the results: per task, by adding up the scores on the four tasks and by the motor quotient, calculated by adding up the scores.

METHOD

The use of the motor batteries in different countries stressed the need to confront, by means of a systematic review, information on the psychometric elements taken into account in the cross-cultural validation of the TGMD-2, MABC-2 and KTK. Thus, an exploratory study was undertaken and, in this kind of reviews, the documents studied are selected through a systematic method, based on the separation per themes and research axes.

To comply with the criteria of the Prisma Declaration, (Urrutia & Bonfill, 2010), two evaluators independently looked for articles published in English and Portuguese, without time restrictions, in six databases: Web of Science, Science Direct, Lilacs, Scopus, Pubmed and The Scientific Electronic Library Online – SciELO. The keywords used were: “MABC”, “TGMD” and “KTK”, all of which were combined, using the Boolean operator “and”, with the descriptors “validity” and “validade”. First, the analyses were based on the titles of the articles, identifying the studies that could possibly attend to the study objective. Then, the articles were excluded whose abstracts did not fit. Finally, the full articles were analyzed, using articles and abstracts that were unavailable, repeated or did not fit the objectives as an exclusion criterion. The searches were undertaken between September and October 2014.

The inclusion criterion only considered scientific articles for the cross-cultural validation of the motor batteries KTK, MABC-2 and TGMD-2, excluding the following from the analysis: (a) books, book chapters, dissertations and theses; (b) unavailable articles and abstracts; (c)

repeated articles and abstracts; (d) lectures and congress abstracts.

Based on this review, we elaborated an epistemological reflection on the criteria used to validate these instruments, with a view to studying the scientific knowledge from a critical viewpoint, reflecting and discussing on the methods, cultural significance, place, limits and emergence of knowledge (Aranha & Martins, 2003).

RESULTS

In total, 734 articles were found, of which only 10 studies were maintained after the exclusion criteria. Figure 1.

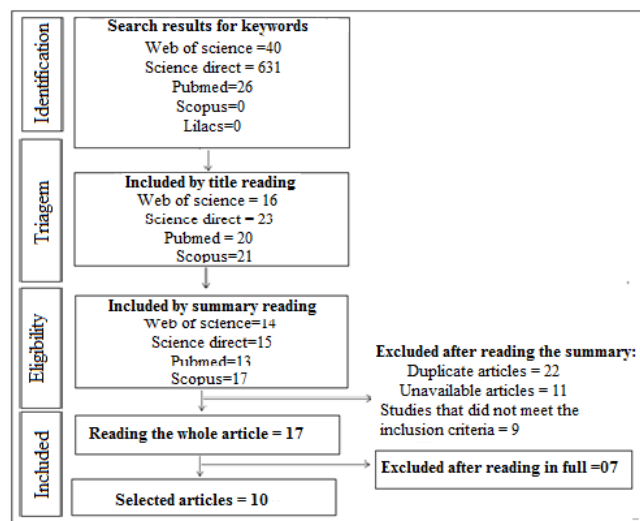


Figure 1. Flowchart of articles found

Five articles were found related to the motor battery TGMD-2, four of which present all psychometric test elements in terms of validity and reliability. In two studies, however, the lack of applicability was observed concerning the criterion validity and the inter/intra-rater reliability (Table 1).

What the MABC-2 battery is concerned, five studies were selected, only two of which presented all validity and reliability criteria. As for the remainder, mainly the content validity criterion was lacking (Table 2).

Regarding the KTK motor battery, no articles were found with the proper inclusion criteria proposed in this study.

Table 1

Description of the applicability in relation to the validity and reliability criteria of the articles selected for the motor battery Test of gross motor development- TGMD-2.

Authors	Main Objective	Age, sample, country and year	Main results	Validity			Reliability		
				Content	Construct	Criterion/competitor/convergent	Retest test	Internal consistency	Inter / Intra Reviewer
Kim, Han & Park	Assess the validity and reliability of the TGMD 2 in Korean children.	The study participants were 121 children between 5 and 6 years old. Korea. 2014.	The research showed adequate reliability and validity of the TGMD-2 for Korean preschool children. In addition, the TGMD-2 is appropriate to assess the motor development of Korean children.	Applied	Applied	Applied	Applied	Applied	Applied
Valentini	Translate the TGMD-2 to Portuguese and validate the test in Brazil.	Participants were 2,674 children between 3 and 10 years old from the five Brazilian regions. 2012.	The motor battery TGMD-2 is a valid and reliable instrument for application in Brazilian children.	Applied	Applied	Applied	Applied	Applied	Applied
Hartman, Jonker, Visscher	Analyze the psychometric properties of the TGMD-2 in children with visual impairments (VI)	The participants were 75 visually impaired children between 6 and 12 years of age. The Netherlands. 2010.	The TGMD-2 battery is an adequate instrument to assess the motor skills of visually impaired children of primary school age.	Applied	Applied	Applied	Applied	Applied	Applied
Simons et al.	Assess the validity and reliability of the TGMD-2 in intellectually disabled Flemish children.	The participants were 99 children between 7 and 10 years of age. 2008.	The TGMD-2 battery is adequate for the motor assessment of children with mild mental disability.	Applied	Applied	Applied	Applied	Applied	Applied
Seonijin et al.	Investigate the reliability and validity of the (TGMD-2) for South Korean children	The participants were 121 Korean children (5 and 6 years). South Korea. 2014	The TGMD-2 is an adequate tool to assess the motor skills in South Korean children.	Applied	Applied	Not Applied	Applied	Applied	Not Applied

Table 2

Description of the applicability in relation to the validity and reliability criteria of the articles selected for the motor battery Movement Assessment Battery for Children – MABC-2.

Authors	Main Objective	Age, sample, country and year	Main results	Validity			Reliability		
				Content	Construct	Criterion/competitor/convergent	Retest test	Internal consistency	Inter / Intra Reviewer
Valentini, Ramalho & Oliveira	Translate the MABC-2 to Brazilian Portuguese; Examine the reliability and validity of the translated MABC-2.	The participants were 844 children between 3 and 13 years of age. Brazil. 2014	This study confirmed that the original standardized scores established for the MABC-2 are valid for applicability in children from the South of Brazil	Applied	Applied	Applied	Applied	Applied	Applied
Schoemaker, Smits-Engelsman, & Bjongmans	Investigate the psychometric properties of the MABC checklist and its utility as a screening tool	The participants were 120 children between 6 and 11 years of age. The Netherlands. 2012.	The MABC-2 battery is valid for all proposed age bands of Dutch children, except for children aged 8 years old, in whom the battery did not detect many motor problems.	Not applied	Applied	Applied	Not applied	Applied	Not applied
Hua, Gu, Meng, & Wu	Assess the validity and reliability of age band 1 of the MABC-2 to prepare for its standardization in mainland China	The participants were 1823 children between 36 and 72 months of age. China. 2013.	To apply the MABC-2 battery in Chinese children, for age band 1, adjustments in the items are needed.	Applied	Applied	Applied	Applied	Applied	Applied
Ellinoudis, Evaggelino, Kourtessis, Konstantinidou, Venetsanou, Kambas	Examine specific aspects of the reliability and validity of age band 1 of the MABC-2	The participants were 183 children between 36-64 months of age. Northern Greece.	These study results suggest that the MABC-2 is a valid and reliable tool for children in Northern Greece.	Not applied	Applied	Not applied	Applied	Applied	Applied
Schoemaker, Niemeijer, Flapper, Smits-Engelsman	Investigate the validity and reliability of the Movement Assessment Battery for Children-2 Checklist (MABC-2)	The participants were 383 children between 5 and 9 years of age. 2012. The Netherlands and Belgium.	The MABC-2 battery complies with the validity and reliability standards for children from the Netherlands and Belgium.	Not applied	Applied	Applied	Not applied	Applied	Not applied

DISCUSSION

The intent in this study was to confront information from the literature about the psychometric elements taken into account in the cross-cultural validation of the TGMD-2, MABC-2 and KTK batteries in different countries, by means of a systematic review; as well as to elaborate an epistemological reflection on the criteria used in the elaboration and validation of these tools. According to the research findings, five cross-cultural validation articles were found for the motor battery TGMD-2, five for the MABC-2 and none for the KTK. Concerning the main results of the validation articles identified in this study, they were similar in that all authors concluded that the motor batteries were valid and reliable in the selected samples. Only some reservations were found in the studies by Schoemaker, Smits-Engelsman, and Bjonngmans, (2003), in which the MABC-2 battery did not detect children with motor problems in the age range of eight years in the Netherlands; and in the study by Hua et al. (2013), in which the authors highlighted the need for adjustments in the items of the MABC-2 battery for Chinese children in a given age range.

Among the 10 articles selected, 6 presented all validity and reliability criteria analyzed in this study (Hartman, Jonker, & Visscher, 2010; Hua et al., 2013; Joo & Clark, 2014; Kim, Han, & Park, 2014; Simons et al., 2008; Valentini, 2012). The other four articles, however, presented incongruences with regard to the criteria adopted. It was verified that, in the study by Ellinoudis et al., (2011) the content analysis for the validity of the battery was not identified. In Seonijin et al. (2014), the criterion/concurrent/convergent analysis was not identified for the validity and the inter/intra-rater analysis for the reliability. In the studies by Schoemaker et al. (2003) and Schoemaker et al., (2012) the content analysis for validity and the test-retest and inter/intra-rater analysis were not identified for reliability.

Hence, it is observed that there is no unanimity on the use of the recommended criteria (Borsa, Damásio, & Bandeira, 2012; Pasquali, 2006) for the cross-cultural validation of the instruments under analysis. These findings

are in line with the study by Cools, de Martelaer, Samaey and Andries (2009), in which the authors developed a study on the psychometric elements used in the creation of seven motor assessment instruments and also verified that the authors did not use the same principles or psychometric criteria to validate the batteries.

The adoption of different criteria in the cross-cultural validations appoints weaknesses of the motor batteries in certain populations, directly interfering in the results of the scientific productions. Wright, Sugden, and Tan (1994) highlight that, although satisfactory results were found in the application of a certain motor battery, it would need some adjustments to cover particularities of children from other contexts. That was also verified in the study by Miyahara et al., (1998) developed in Japan, in which the authors highlighted the need to adequate the tasks in the test to the experiences of the children in that environment. Rosblad and Gard (1998) also evidenced differences between the American and Swedish samples in the performance of the skills tests using ball and balance. Chow, Henderson, and Barbett, (2001) observed that the performance of children from Hong Kong was superior to that of American children on the balance and manual skills tests, attributing these differences to the particularities of the environment they lived in. In addition, Smits-Engelsman, Henderson, and Michels (1998) related two motor batteries and found differences in the tests' capacity, attributing these differences to the particularities of the environment each test was constructed in. The application of the motor batteries in different countries faced other difficulties, such as high asymmetry and social inequality rates, motivational aspects, mechanical and anthropometric aspects, familiarity or not with a certain motor task and different meanings the individual attributed to that task (Payne & Isaacs, 2007)

This lack of coherence in the psychometric procedures or criteria the researchers use in the construction and validation of the motor assessment instruments makes it difficult to understand the process as a whole (Libardoni dos Santos, 2014). The International Test Commission (ITC, 2010) had already highlighted

that many of the measures used in other countries do not follow all recommended procedures for validation in other populations, proposing guidelines to bridge these divergences. In this perspective, Cools et al., (2009) highlight that the most fundamental criticism concerning the validation of motor assessment instruments is that they do not have the same psychometric quality as the instruments used in other areas, like psychology for example.

Also from a general perspective, when confronting the psychometric procedures appointed in Pasquali, (2006) with the psychometric procedures used in the motor area (Pictures 1 and 2), some weaknesses are common. The motor instruments tend to follow the strict caution taken with elements like validity and reliability, but do not deepen these elements as a whole in accordance with psychometric recommendations (Borsa et al., 2012; Libardoni dos Santos, 2014), mainly when considering validity (Table 1 and 2).

When analyzing each battery separately, it could be observed in the articles about the TGMD-2 battery that, among the five articles found (Picture 2), four presented all psychometric elements needed to validate and/or adapt a motor assessment instrument (Kim et al., 2014; Simons et al., 2008; N.C Valentini, 2012). In a study by Seonijin et al., (2014), however, the lack of applicability was observed with regard to the criterion validity and inter/intra-rater reliability. The importance of caution with the lack of testing of these elements is highlighted, as the criterion validity is intended to verify if the instrument can identify the persons who are truly better at performing a given skill (Viana, 1989).

Hence, the (predictive or concurrent) criterion validity of a test can suffer the influence of other factors not associated with the test, which can affect the magnitude of the validity coefficient. In addition, the lack of criterion validity does not always reflect the lack of validity, but can indicate the possibility that the criterion is totally questionable (Libardoni dos Santos, 2014). As regards the inter/intra-rater reliability, it is important to highlight that not testing this element can entailing measuring errors, as it indicates the extent to which the differences in

the scores derive from variations in the characteristic examined and not from casual errors by the raters (Raymundo, 2009).

Concerning the MABC-2 battery, five studies were selected, only two of which (Hua et al., 2013; Valentini, Ramalho, & Oliveira, 2014) presented all recommended validity and reliability criteria (Borsa et al., 2012; Pasquali, 2006). In the remainder, mainly the content validity criterion was missing. Tritschler (2003) highlights the importance of content validity, as validity is a fundamental element in the elaboration of an assessment tool. According to Sim and Arnell, (1993), validity refers to the relation between the measure and the previously defined phenomenon (construct), that is, what the instrument intends to measure in a specific context. Validity is considered as an inherent part not only of the instrument, but as an attribute of the collection and analysis, that is, of the test as a whole (Morrow, et al., 1995).

Concerning the KTK battery, no articles were found with the same inclusion criteria as in this study, although the instrument is frequently used in Brazil. That can be due to the uncommon validation of this motor assessment instrument in other countries, and even to the unavailable studies in the selected databases.

It is important to highlight that validations and cross-cultural adaptations are complex tasks that require planning and rigor (Cassepp-Borges, Balbinotti, & Teodoro, 2010), in which the validity of the tool needs to be proven in another cultural context (Borsa et al., 2012). The lack of criterion can produce maladapted instrument with incoherent or hardly reliable data (Borsa et al., 2012). An assessment instrument that is not valid ends up losing its essence (Cools et al., 2009).

Thus, based on all questions raised, inquiries emerge on the existing paradigms in Physical Education, and consequently on the motor behavior that is part of it. These paradigms include the lack of dialogue, whether between Physical Education and other areas or among motor behavior researchers themselves. Until the establishment of a dialogue on the standardization of instrument validations and constructions, breaking with that paradigm, a

long road lies ahead, facing great resistance, which can be called an epistemological obstacle (Bachelard, 1996).

These obstacles have made it difficult to improve the instruments used in the area and to reach a consensus on the psychometric elements that should be used in the validations, resulting in studies with less reliable data, which are incapable of reflecting the reality or of being compared and reapplied in different cultures. These limitations could be overcome in other studies if Physical Education researchers were willing to dialogue in order to find the best answers together with the researchers who have been developing similar studies and with more positive results. This lack of interaction in the area limits the psychometric knowledge, but also the understanding that the phenomenon can reach beyond the limited world it was created for. Therefore, its applicability needs to be understood and verified in other contexts and perhaps further explored by other areas of science.

Similarly, this lack of interaction among the researchers in the area was clear, in view of the different existing motor batteries that, overall, aim to assess similar criteria but do not find a consensus. Therefore, equivalent knowledge needs to be constructed on the validations of motor instruments, the interactions with the world and among the researchers (Oliveira & Caminha, 2014), but mainly on the understanding of their objectives, goals and targets, with a view to a better understanding of the area (Werner, 1998).

These findings indicate that, in recent years, instead of trying to improve existing tools, the researchers end up putting forward their truths by creating new instruments. Although these have been important research tools, they were insufficient to solve the shortages found in the area. In addition, the continued transmission of this knowledge without proper questioning is highlighted as, according to Fensterseifer (2009), the problem is not about “transmitting truths”, but about not revealing the nature of these truths,

that is, assuming and presenting the limitations implicit in these research instruments for example.

Nevertheless, the standardization in general is going through a profound crisis (Fensterseifer, 2009). That can be beneficial in this area, considering that, at times of crisis, if the established values are no longer sustained, there is space to affirm new values. Hence, the knowledge validation process is open, in which the human truths assume their precarious status and are always suspected. They only survive as long as they are “co-validated” by concrete subjects in equally concrete situations, which would consequently turn the dialogue into an almost compulsory tool in knowledge construction.

CONCLUSION

Although some studies presented similar validation criteria, there is no consensus on the use of the recommended criteria for the cross-cultural validation of the instruments in question. The authors diverge on the psychometric elements taken into account in the cross-cultural validations and there was no consensus on the validation criteria of all studies found. Although there are no regulations among the researchers in the area, the psychometric elements (validity and reliability) should be respected as a whole, in the cross-cultural validations as well as in the creation of motor assessment tools, thus achieving more valid and reliable results.

As study limitations, we highlight the sole use of articles published in English and Portuguese, which may have resulted in the exclusion of other articles about the theme that were published in other languages. The same limitation extends to the databases used, so that these results cannot be generalized to all publications in other databases. Finally, it is highlighted that some articles were not considered in the analysis because they were not accessible to the authors, which may have limited the interpretation of the results.

Acknowledgments:

Nothing to declare

Conflict of interest:

Nothing to declare

Funding:

Nothing to declare

REFERENCES

- Aranha, M. L. d. A., & Martins, M. H. P. (2003). *Filosofando: introdução à filosofia* (Vol. 3). São Paulo: Moderna.
- Bachelard, G. (1996). *A formação do espírito científico: contribuição para uma psicanálise do conhecimento*. Rio de Janeiro: Contraponto.
- Barnett, A., Henderson, S., & Sugden, D. (2007). *Movement Assessment Battery for Children – MABC2*. London: Psychological Corporation.
- Berleze, A., Haefner, L. S. B., & Valentini, N. C. (2007). Desempenho motor de crianças obesas: uma investigação do processo e produto de habilidades motoras fundamentais. *Revista Brasileira de Cineantropometria e desempenho Humano*, 9(2), 134-144.
- Borsa, J. C., Damásio, B. F., & Bandeira, D. R. (2012). Adaptação e Validação de Instrumentos Psicológicos entre Culturas: Algumas Considerações. *Paidéia*, 22(53), 423-432.
- Cassepp-Borges, V., Balbinotti, M. A. A., & Teodoro, M. L. M. (2010). *Tradução e validação de conteúdo: Uma proposta para a adaptação de instrumentos*. In L. Pasquali (Org.), *Instrumentação psicológica: Fundamentos e prática*. Porto Alegre: Artmed.
- Chow, S. M. K., Henderson, S. E., & Barnett, A. L. (2001). The movement assessment battery for children: a comparison of 4-year-old to 6-year-old children from Hong Kong and de United States. *The American Journal of Occupational Therapy* 55, 55-61.
- Cicchetti, D. V., & Rourke, B. P. (2004). *Methodological and biostatistical foundations of clinical neuropsychology and medical and health disciplines*: CRC Press.
- Cools, W., De Martelaer, K., Samaey, C., & Andries, C. (2009). Movement skill assessment of typically developing preschool children: A review of seven movement skill assessment tools. *Journal of Sports Science and Medicine*, 8, 154-168.
- Ellinoudis, T., Evaggelinou, C., Kourtessis, T., Konstantinidou, Z., Venetsanou, F., & Kambas, A. (2011). Reliability and validity of age band 1 of the Movement Assessment Battery for Children - Second Edition. *Research in Developmental Disabilities*, 32(3), 1046-1051. doi:10.1016/j.ridd.2011.01.035
- Fensterseifer, P. V. (2009). Epistemologia pratica e pedagógica. *Rev. Bras. Cienc. Esporte*, 30(3), 203-214.
- Freitas, D. L., Lausen, B., Maia, J. A., Gouveia, E. R., Thomis, M., Lefevre, J., Malina, R. M. (2016). Skeletal Maturation, body size, and motor coordination in youth 11-14 years. *Medicine & Science in Sports & Exercise*, 48(6), 1129-1135
- 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(2), 143-159
- Henderson, S. E., Sugden, D. A., & Barnett, A. L. (2007). *Movement Assessment Battery for children – 2 Examiner’s Manual*. London: Harcourt Assessment,.
- Hua, J., Gu, G., Meng, W., & Wu, Z. (2013). Age band 1 of the Movement Assessment Battery for Children-Second Edition: Exploring its usefulness in mainland China. *Research in Developmental Disabilities*, 34(2), 801-808. doi:10.1016/j.ridd.2012.10.012
- Joo, K. & Clark, J. E. (2014). Validity and Reliability of the TGMD-2 for South Korean Children. *Journal of Motor Behavior*, 46(5), 351-356.
- Kim, C. L., Han, W. H., & Park, H. (2014). Reliability and Validity of the Test of Gross Motor Development-II in Korean Preschool Children: Applying AHP. *Research in Developmental Disabilities*, 35(4), 800-807.
- Kiphard, E. J., & Schilling, V. F. (1974). *Körperkoordinations-test für Kinder – KTK*, Beltz Test. Weinheim: Gmbh.
- Krebs, R. J., Duarte, M. G., Nobre, G. C., Nazario, P. F., & Libardoni dos Santos, J. O. (2011). Relação entre escores de desempenho motor e aptidão física em crianças com idades entre 07 e 08 anos. *Revista Brasileira de Cineantropometria e Desempenho Humano*, 13(2), 94-99.
- Libardoni dos Santos, J. O. (2014). *Aspectos da validade de conteúdo e construto de tarefas motoras* (Dissertação de Doutorado). Universidade do Estado de Santa Catarina, Florianópolis.
- Liu, T., Hamilton, M., & Smith, S. (2015). Motor proficiency of the Head Start and typically developing children on MABC-2. *Journal of Child and Adolescent Behaviour*, 3(2), 1-4.
- Luz, O. G. L., Seabra, A. F. T., Santos, R., Padez, C., Ferreira, J. P., & Silva, M. J. C. (2015). Associação entre IMC e teste de coordenação corporal para crianças (KTK). Uma meta-análise. *Revista brasileira de medicina do esporte*, 21(3), 230-235.
- Martins, V., Silva, A. J., Marinho, D. A., & Costa, A. M. (2015). Desenvolvimento motor global de crianças do 1º ciclo do ensino básico com e sem prática prévia de natação em contexto escolar. *Motricidade*, 11(1), 87-97.
- Miyahara, M., Tsujii, M., Hanai, T., Jongmans, M., Barnett, A., Henderson, S. E., & Kageyama, H. (1998). The movement Assessment Battery for Children: A preliminary investigation of its

- usefulness in Japan. *Human Movement Science*, 17(4), 679-679.
- Moreira, R. S., Magalhães, L. C., Dourado, J. S., Lemos, S. M., & Alves, C. R. (2014). Factors influencing the motor development of prematurely born school-aged children in Brazil. *Research Development Disabilities*, 35(9), 1941-1951.
- Oliveira, G. M., & Caminha, O. I. (2014). Epistemologia genética e educação física: algumas implicações pedagógicas. *Revista Brasileira de Psicologia Escolar e Educacional*, 18(1), 57-65.
- Pasquali, L. (2006). *Técnicas de Exame Psicológico: manual* (Vol. 2). São Paulo: Caso do Psicólogo.
- Pasquali, L. (2010). *Instrumentação Psicológica*. Porto Alegre: Artmed.
- Payne, V. G., & Isaacs, L. D. (2007). *Desenvolvimento Motor Humano: Uma Abordagem Vitalícia*. Rio de Janeiro: Guanabara Koogan.
- Piek, J., P, Gasson, N., & Barrett, N. (2002). Limb and gender differences in the development of coordination in early infancy. *Human Movement Science*, 21(5), 621-639.
- Raymundo, V. P. (2009). Construção e validação de instrumentos: um desafio para a psicolinguística. *Letras de Hoje*, 44(3), 86-93.
- Reichenheim, M. E., & Moraes, C. L. (2007). Operacionalização de adaptação transcultural de instrumentos de aferição usados em epidemiologia. *Revista de Saúde Pública*, 41(4), 665-673.
- Ribeiro, A. P., Trombini-Souza, F., Iunes, D. H., & Monte-Rasso, V. V. (2006). Confiabilidade inter e intraexaminador da fotopodometria e intra-examinador da fotopodoscopia. *Revista Brasileira de Fisioterapia*, 10(4), 435-439.
- Rosa Neto, F. (2002). *Manual de avaliação motora* (Vol. 1). Porto Alegre: Artmed - Artes Médicas.
- Rosblad, B., & Gard, L. (1998). The assessment of child with developmental coordination disorders in Sweden: a preliminary investigation of the Movement ABC. *Human Movement Science*, 17, 711-719.
- Rudd, J., Butson, M., Barnet, L., Farrow, D., Berry, J., Borkoles, E., & Polman, R. (2015). Physical literacy development in Australian youth: A current concern: How should we measure movement competence in children? *Journal of Science and Medicine in Sport*, 19(1), e62.
- Schoemaker, M. M., Niemeijer, A. S., Flapper, B. C. T., & Smits-Engelsman, B. C. M. (2012). Validity and reliability of the Movement Assessment Battery for Children-2 Checklist for children with and without motor impairments. *Developmental Medicine and Child Neurology*, 54(4), 368-375. doi:10.1111/j.1469-8749.2012.04226.x
- Schoemaker, M. M., Smits-Engelsman, & Bjonngmans, M. J. (2003). Psychometric properties of the Movement Assessment Battery for Children-Checklist as a screening instrument for children with a developmental co-ordination disorder. *British Journal of Educational Psychology*, 73(3), 425-441.
- Seonijin, K., Min Joo, K., Valentini, N. C., & Clark, J. E. (2014). Validity and Reliability of the TGMD-2 for South Korean Children. *Journal Motor Behavior*, 56(5), 351-356.
- 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.
- Smits-Engelsman, B. C. M., Henderson, S. E., & Michels, C. G. J. (1998). The assessment of children with developmental coordination disorders in the Netherlands : the relationship between the movement assessment battery for children and the korperkoordinations test fur kinder. *Human Movement Science*, 17, 699-709.
- Souza, C., Ferreira, L., Catuzo, M. T., & Correa, U. C. (2007). O teste ABC do movimento em crianças de ambientes diferentes. *Revista Portuguesa de Ciências do Desporto*, 7(1), 36-47.
- Souza, M. C., Berleza, A., & Valentini, N. C. (2008). Efeitos de um programa de educação pelo esporte no domínio das habilidades motoras fundamentais e especializadas: ênfase na dança. *Revista da Educação Física*, 19(4), 509-519.
- Sun, W., Chou, C.-P., Stacy, A. W., Ma, H., Unger, J., & Gallager, P. (2007). SAS and SPSS macros to calculate standardized Cronbach's alpha using the upper bound of the phi coefficient for dichotomous items. *Behavior Research Methods*, 39, 71-84.
- Tritschler, K. (2003). *Medida e avaliação em educação física e esportes* (5th ed.). São Paulo: Manole.
- Ulrich, D. A. (2000). *TGMD 2 – Test of gross motor development examiner's manual* (2nd ed.). Austin, Texas: Pro-Ed.
- Urru'tiaa, G., & Bonfill, X. (2010). PRISMA declaration: A proposal to improve the publication of systematic reviews and meta-analyses. *Medicina Clínica*, 135(11), 507-511.
- Valentini, N. C. (2012). Validity and reliability of the TGMD-2 for Brazilian children. *Journal of Motor Behavior*, 44(4), 275-280.
- Valentini, N. C., Ramalho, M. N., & Oliveira, M. A. (2014). Movement Assessment Battery for Children-2: Translation, reliability, and validity for Brazilian children. *Research in Developmental Disabilities*, 35(3), 733-740. doi:10.1016/j.ridd.2013.10.028
- Van Keulen, G. E., Benda, R. N., Ugrinowitsch, H., Valentini, N. C., & Krebs, R. J. (2016). A influência de uma intervenção utilizando a prática variada e em blocos no desempenho das habilidade de controle de objetos em crianças eutróficas de 6 e 7 anos. *Revista da Educação Física*, 27(1), 2707-16.
- Viana, H. M. (1989). *Introdução à avaliação educacional*. São Paulo: IBRASA.
- Wright, H., Sugden, D. A., & Tan, J. (1994). Identification of children with movement problems in singapore: usefulness of the

movement ABC checklist. *Adapted Physical Activity Quarterly*, 11(2), 150-57.

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



Todo o conteúdo da revista **Motricidade** está licenciado sob a [Creative Commons](https://creativecommons.org/licenses/by-nc/4.0/), exceto quando especificado em contrário e nos conteúdos retirados de outras fontes bibliográficas.