Motor skills assessments: support for a general motor factor for the Movement Assessment Battery for Children-2 and the Bruininks-Oseretsky Test of Motor Proficiency-2

Avaliação de habilidades motoras: suporte a um fator motor geral para o Movement Assessment Battery for Children-2 e o Bruininks-Oseretsky Test of Motor Proficiency-2

Paola Matiko Martins Okuda,[1] Melissa Pangelinan,[2] Simone A. Capellini,[3] Hugo Cogo-Moreira[1]

Abstract

Objective: To evaluate the construct validity and model-based reliability of general and specific contributions of the subscales of the Movement Assessment Battery for Children-2 (MABC-2) and Bruininks-Oseretsky Test of Motor Proficiency-2 (BOT-2) when evaluating motor skills across a range of psychiatric disorders.

Methods: Confirmatory factor analysis (CFA) and bifactor analysis were conducted on BOT-2 data from 187 elementary school students (grades 1 to 6) (mean age: 113 ± 20 months; boys: n = 117, 62.56%) and on MABC-2 data from 127 elementary school students (grade 1) (mean age: 76 ± 2 months; boys: n = 58, 45.67%).

Results: The results of the CFA fit the data for multidimensionality for the BOT-2 and presented poor fit indices for the MABC-2. For both tests, the bifactor model showed that the reliability of the subscales was poor.

Conclusions: The BOT-2 exhibited factorial validity with a multidimensional structure among the current samples, but the MABC-2 showed poor fit indices, insufficient to confirm its multidimensional structure. For both tests, most of the reliable variance came from a general motor factor (M-factor), therefore the scoring and reporting of subscale scores were not justified for both tests.

Keywords: Bruininks-Oseretsky Test of Motor Proficiency, Second Edition, Movement Assessment Battery for Children, Second Edition, psychometrics, motor skills, evaluation report.

Resumo

Objetivo: Avaliar a validade de construto e a confiabilidade das subescalas do Movement Assessment Battery for Children-2 (MABC-2) e do Bruininks-Oseretsky Test of Motor Proficiency-2 (BOT-2) com base em modelos das contribuições gerais e específicas ao avaliar habilidades motoras em transtornos psiquiátricos.

Métodos: Foram realizadas análise fatorial confirmatória (AFC) e análise bifatorial em dados de 187 escolares do primeiro ao sexto ano do ensino fundamental (idade média: 113 ± 20 meses; meninos: n = 117, 62,56%) que foram avaliados com o BOT-2, e em dados de 127 escolares do primeiro ano do ensino fundamental (idade média: 76 ± 2 meses; meninos: n = 58, 45,67%) avaliados com o MABC-2.

Resultados: Os resultados da AFC apresentaram índices de ajuste satisfatórios de multidimensionalidade para o BOT-2 e apresentaram índices de ajuste insatisfatórios para o MABC-2. Para ambos os testes, o modelo bifatorial mostrou que a confiabilidade das subescalas era ruim.

Conclusão: O BOT-2 apresentou validade fatorial com uma estrutura multidimensional entre as amostras utilizadas, mas o MABC-2 apresentou índices de ajuste insatisfatórios, insuficientes para confirmar sua estrutura multidimensional. Para ambos os testes, a maior parte da variância confiável veio de um fator motor geral (fator-M), portanto, a pontuação e o relato dos escores das subescalas não se justificaram para ambos os testes.

Descritores: Bruininks-Oseretsky Test of Motor Proficiency, Second Edition, Movement Assessment Battery for Children, Second Edition, psicometria, habilidades motoras, relatório de avaliação.

Introduction

Motor skills serve as the basis not only for sports and recreation, but are embedded in all activities of daily living. The identification of movement difficulties in children is crucial to understanding the biological basis of neurodevelopmental disorders, such as developmental coordination disorder[1] and neurological soft signs.[2] Recently, as signs of motor dysfunction are evidenced across a range of psychiatric disorders, especially schizophrenia, the Research Domain Criteria Initiative (R-DoC) propose a domain of motor systems in an attempt to understand and explain the relations between motor circuits and the pathophysiology of psychiatric disorders.[3] Hence, motor skills assessment is fundamental for identifying and understanding the pathophysiology underlying neurodevelopmental and psychiatric disorders and for implementing early intervention and effective rehabilitation treatment plans and verifying the potential relationship between them.

While there is no gold standard to measure children's motor abilities, the Movement Assessment Battery for Children, Second Edition (MABC-2)[4] and the Bruininks-Oseretsky Test of Motor Proficiency (BOT-2)[5] are the tools most commonly used in both clinical and research settings. The former assesses three dimensions: manual dexterity, aiming and catching, and balance. It comprises eight items/tasks and has had its reliability and construct validity demonstrated in European and Asian countries.[6-9] Although the MABC-2 has been assessed for validity in the Brazilian population,[10] construct validity was not contemplated. The BOT-2 assesses four dimensions: fine manual control, manual coordination, body coordination, and strength and agility. It also comprises eight items/tasks. Although construct validity has not been assessed by external researchers, factorial validity is provided in the assessment manual,[5] with good fit statistics that provide validity evidence for the four motor-areas (Table 6.10 in the manual).

The conceptual model underlying the items of both tools is multidimensional; however, being a multidimensional construct per se does not guarantee that each subscale is reliable (how well a latent variable is represented by a given set of items [i.e., the quality of its indicators]), replicable across studies,[11] or viable (capable to be sustained in the model); psychometric features of multidimensional constructs are evaluated through indices derived from a bivariate model. These indices are useful to describe (a) the quality of unit-weighted total and subscale score composites, and (b) the specification and quality of a measurement model in structural equation modeling.[11]

There is only one study in the area of motor assessment[7] that suggests a common "general motor ability" construct underlying the subscales. General motor ability comprehends a general factor that underlies the subscales within a battery/test, which may influence the performance of the subscales, crossing over all the items effectively and capturing their shared content with a unifying concept, whereas the specific factors (subscales) account for response variation that is unique or particular to item subsets.[12]

Formal procedures to evaluate the reliability and viability of the subscales in the presence of a general motor ability factor were not previously tested for MABC-2 and BOT-2 and are conducted via bifactor modeling. Clinically, it is fundamental to determine if the variance (i.e., information) captured by the motor subtest is reliable and viable when controlled by a general motor ability, as such information has a direct effect on how motor assessment scores are conducted (justification for the scoring and its reporting) and subscales interpreted.[11]

The formal procedure, which enables the investigation of the psychometric features of specific factor in the presence of a general factor, is the bifactor model (also known as the nested factors/direct hierarchical/general-specific model).[13] Bifactor models are a type of specification of confirmatory factor models.[14] Regarding adequacy, bifactor models are less restrictive due to more free parameters[15] and consequently they will have better fit indices when compared to other commonly more restrictive multidimensional solutions, such as correlated-factor models or second order models.[11,13]

Consequently, we aimed to answer the following questions: 1) Are the subscales for MABC-2 and BOT-2, the tools most commonly used for motor assessment, valid and reliable? 2) Is the scoring and reporting of subscale scores justified? 3) If justified, how much reliable variance do the subscale scores provide after controlling for a general motor factor (M-factor)? In sum, the overall objective was to evaluate the model-based reliability of general and specific contributions of the subscales of the BOT-2 and MABC-2, thus providing insights regarding the construct validity (factorial validity or validity based on internal structure) for both tools.

Method

This research was approved by the research ethics committee of Universidade Federal de São Paulo (UNIFESP; protocol 1.325.805). The study participants provided written informed consent, and the study

complied with the Declaration of Helsinki of the World Medical Association.

Participants

The study used previously collected data (convenience sample) from various centers and populations to conduct the analysis involving bifactor modeling. Consequently, we considered at least 10 participants per observed indicator variable as a rule of thumb for a lower bound on adequate sample size,[16] totaling at least 80 children for each instrument. Also, we considered all the available sample (with or without some diagnosis for the BOT-2), because latent trait (construct) models are measured at item level and provide sample-free measurement.[16,17]

Sample for the BOT-2

The BOT-2 was used in a sample of 187 public elementary school students from grades 1 to 6 in São Paulo (mean age: 113 ± 20 months; boys: n = 117, 62.56%). Students with learning difficulties (n = 20; 10.7%) and learning disorders (dyslexia: n = 20, 10.7%; language-based learning disability: n = 20, 10.7%) were included in this sample.

All students were assessed with the full version of the BOT-2 in single 50 to 60-minute sessions, applied by a trained occupational therapist, in a classroom or courtyard provided by the school or in the attendance rooms of the Center of Education and Health Studies (Centro de Estudos em Educação e Saúde – CEES), Universidade Estadual Paulista (UNESP), Marília, and at the Outpatient Clinic of Child Neurology – Learning Disorders, at Hospital das Clínicas da Faculdade de Medicina de Botucatu, UNESP. A detailed description of the diagnostic criteria used for the sample can be found in the online-only supplementary material, Appendix 1.

Sample for the MABC-2

For the MABC-2,[4] we used a new sample of children from São Paulo: 127 public elementary school students from grade one (mean age: 76 ± 2 months; boys: n = 58, 45.67%). Based on school records and parental reports sent to the school, the participants had no pre-, peri-, or post-natal difficulties, no delay in neuropsychomotor or language development, and no behavioral problems. None of the participants had received a diagnosis of intellectual or physical disability, including pervasive developmental disorder or general medical conditions such as cerebral palsy, muscular dystrophy, or hemiplegia; neither did they have any learning disorders or difficulties. All students were assessed with the MABC-2 in single 30 to 50-minute

sessions, applied by a trained occupational therapist, in a classroom provided by the school.

Procedures

The BOT-2 is an objective instrument widely used in clinical and research settings to measure gross and fine motor functioning for individuals aged 4-21 years.[5] The BOT-2 provides scores in four domains of motor competence and a total motor composite score, which includes all four domains: 1) fine manual control: fine motor precision and integration; 2) manual coordination: manual dexterity and upper-limb coordination; 3) body coordination: bilateral coordination and balance; and 4) strength and agility: running speed and agility, and strength.

Scoring is based on the results of goal directed-activities, where the total score in each item is converted to a scale score for each item; then, the pairs of items that form the domains are converted to a standard score, and the sum of the domain standard scores is converted to a total composite score. All these scores were considered continuous variables.

The MABC-2 is also an objective instrument widely used in clinical and research settings to measure gross and fine motor skills with normative data for three age bands. For this study, only age band 1 (3 years to 6 years and 11 months) was used given the participants' ages. The MABC-2 provides scores in three domains of motor competence and a total motor score, which includes all three domains named as following (as described in the manual): 1) manual dexterity: posting coins with one's preferred hand, posting coins with one's non-preferred hand, threading beads, and drawing a trail; 2) aiming and catching: catching beanbag and throwing beanbag onto mat; and 3) balance: one-leg balance best leg, one-leg balance other leg, walking heels raised, and jumping on mats.

Scoring is based on the results of goal directed-activities and errors, where the raw score for each item is converted to a standard score for the items; then, the pairs of items that devise the three domains are converted to a standard score and percentile, and the sum of the domain standard scores is converted to a total motor performance score. All these scores were considered continuous variables.

Data analyses

All statistical analyses were conducted with Mplus 7.4.[18] To verify the dimensional solution of the BOT-2 and MABC-2, a confirmatory factor analysis (CFA) was conducted. Robust maximum-likelihood estimation was used.[18] The following fit indices were used to evaluate the model fit for CFA when all observed variables were

continuous: chi-square, comparative fit index (CFI), Tucker-Lewis index (TLI), and root mean square error of approximation (RMSEA). To demonstrate a good fit to the data, an estimated model should have an RMSEA near or below 0.06 and a CFI and TLI near or above 0.95.[19]

The viability and reliability of BOT-2 and MABC-2 subscales were evaluated using a bifactor model.[14,20] The BOT-2 model consisted of four specific factors, while the MABC-2 model consisted of three specific factors. In relation to CFA, where factors are all free to correlate among each other, the bifactor model should be an orthogonal model, where the relationship among the specific factors themselves and with the general factor is fixed to zero (i.e., no correlation between the factors).[21]

Bifactor models have several potential advantages when compared to other forms of specifying a confirmatory model, particularly when researchers are interested in the predictive relationships between domain-specific factors over and above the general factor. A bifactor model can be used as a less restricted baseline model. It can be used to study the role of domain-specific factors that are independent of the general factor and directly examine the strength of the relationship between the domain-specific factors and their associated items, because the relationship is reflected in the factor loadings. Bifactor models can be particularly useful in testing whether a subset of the domain-specific factors predicts external variables, over and above the general factor, since the domain-specific factors are directly represented as independent factors. Finally, the bifactor model also provides output that paints a more direct interpretation of factor loadings for M-factor vs. the specific factor.[13,17]

The following indices were used to assess the viability of the BOT-2 and MABC-2 subscales: a) coefficient omega (ω),[14,22,23] which is a reliability estimate based on factorial model that estimates the proportion of the observed variance in the total score attributed to all sources of common variances; b) coefficient omega hierarchical (ωh),[20,24] which is a reliability index that judges the degree to which composite scale scores are interpretable as a measure of a single common factor (the coefficient omega hierarchical is computed by dividing the squared sum of the factor loadings on the general factor [model estimated] by the variance of total scores]); c) coefficient omega subscale (ωs),[20,25] which is the percentage of the subscale score variance attributable to a specific group factor of items after removing the reliable variance due to the general factor, i.e., an index reflecting the reliability of a subscale score after controlling for the variance due to the general

factor; and d) explained common variance (ECV), which is the percentage of common variance explained by the general factor, i.e., a type of unidimensionality index directly related to the relative strength of the general factor. It can be defined as the ratio between the explained variance by the general factor and the variance of the general and specific factors. Details of these calculations can be found elsewhere.[20,24]

Coefficient omega (ω), coefficient omega hierarchical (ωh), and coefficient omega subscale (ωs) scores > 0.8 indicate a strong relationship between the latent variable and item scores. An ECV > 0.70 indicates that the instruments should be treated as essentially unidimensional; correspondingly, single common factors were specified.[24]

Results

Descriptive results for the BOT-2 and MABC-2 are available in the online-only supplementary material, Tables S1 and S2. The complete data and the computing code (outputs) can be requested from the corresponding author.

Bruininks-Oseretsky Test of Motor Proficiency, Second Edition (BOT-2)

For the BOT-2, excellent fit indices were found using the CFA. Figure 1 shows the diagram of the underlying model of the BOT-2 tasks (treated as continuous variables), their factor loadings, and the standard errors of loadings (in parentheses). The fit indices for the CFA with four dimensions were: $\chi^2(14) = 20.937$, p = 0.1135; CFI = 0.988; TLI = 0.976; RMSEA = 0.050 (90% confidence interval [90%CI] = 0.000 to 0.093).

The correlations between manual coordination and body coordination, manual coordination and strength and agility, and body coordination and strength and agility were above of 0.90.

Since the four factors were correlated in the BOT-2 model, the bifactor model was applied, where the M-factor was related to all items and the four domains (b1, b2, b3, b4) with the specific items, thus representing specific factors. Considering the bifactor model for BOT-2, the model presented reasonable fit indices: $\chi^2(17) = 38.545$, p = 0.0021; CFI = 0.962; TLI = 0.938; RMSEA = 0.082 (90%CI = 0.048 to 0.117). We then examined the reliability and viability of the four specific factors and found the following coefficient omega subscale scores: fine manual control (b1), ωs = 0.047; manual coordination (b2), ωs = 0.038; body coordination (b3), ωs = 0.014; and strength and agility (b4), ωs = 0.327. Therefore, the coefficients suggested poor viability.

When comparing omega hierarchical ($\omega h = 0.901$) with omega ($\omega = 0.938$) results, we observed that 96.05% of the reliable variance stemmed from the M-factor (0.901/0.938). Therefore, only 3.7% of the variance (0.938-0.901) in the total score could be attributed to multidimensionality caused by four specific factors. The ECV for BOT-2 was 0.83, indicating a strong M-factor.

Figure 2 shows the bifactor model for BOT-2 with their respective standard factor loadings and standard errors. It is important to note that the values of the factor loadings reduce considerably compared to the models shown in Figure 1.

Movement Assessment Battery for Children, Second Edition (MABC-2)

Figure 3 shows the diagram of the underlying model of MABC-2 tasks treated as continuous variables, their factor loadings, and standard errors of the loadings (in parentheses). The fit indices for the CFA with three dimensions were: $\chi^2(32) = 46.569$, $p = 0.0463$; CFI = 0.92; TLI = 0.89; RMSEA = 0.06 (90%CI = 0.008 to

0.095) – results that may be considered reasonable, but not satisfactory.

As observed in Figure 3, four tasks (drawing trail, throwing beanbag onto mat, walking heels raised, and jumping on mats) exhibited factor loadings below 0.3. According to the literature,[26] for an item to remain in the model, the factor loading should be more than 0.3. Despite this, all tasks were kept in the model because, in addition to representing the original model, one of the factors (m2 – aiming and catching) was already at the limit of the minimum number of items (two items) to be considered a latent trait.

Following the same procedure observed for BOT-2, again for the MABC-2, from the three factors, a bifactor model (a general factor and three specific factors) was derived. Although the previous model did not exhibit any good fit indices, the new model specification resulted in excellent fit indices: $\chi^2(26) = 25.560$, $p = 0.4875$; CFI = 1.000; TLI = 1.004; RMSEA = 0.000 (90%CI = 0.000 to 0.069). As evident in Figure 4, there was a reduction in the values of the standard factor loadings of specific
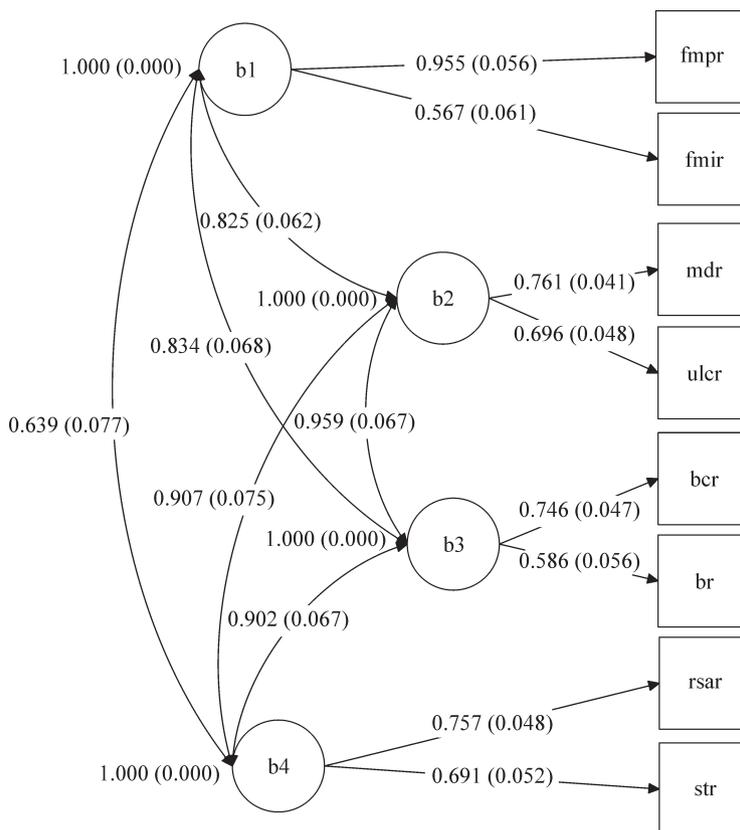
Figure 1 - Confirmatory factor analysis for the Bruininks-Oseretsky Test of Motor Proficiency, Second Edition (BOT-2). b1 = fine manual control; b2 = manual coordination; b3 = body coordination; b4 = strength and agility; fmp = fine motor precision; fmi = fine motor integration; md = manual dexterity; ulc = upper-limb coordination; bc = bilateral coordination; b = balance; rsa = running speed and agility; st = strength; r = raw score.

factors, indicating poor reliability of specific factors and high reliability of the general factor.

The omegas of the subscales were as follows: manual dexterity (m1), ω_s = 0.500; aiming and catching (m2), ω_s = 0.100; and balance (m3), ω_s = 0.243. The relationship between omega hierarchical (ω_h = 0.591) and omega (ω = 0.810) results indicated that 72.96% of the variance came from the M-factor (ω_h/ω = 0.591/0.810) and only 21.9% came from the multidimensionality (ω – ω_h = 0.810-0.591). Moreover, the ECV for the MABC-2 was 0.53, indicating a moderate M- factor.

Discussion

Using the CFA for both BOT-2 and MABC-2, factorial validity with multidimensional structures was found for the studied population. However, large correlations were observed between the BOT-2 factors, indicating that there was no evidence of divergent validity between the factors (the same occurs in the BOT-2 CFA presented in the manual).[5] According to Brown,[17] when the correlation among the factors exceeds 0.85, it may be possible to combine factors to reduce the number of dimensions and thus achieve the most parsimonious set of items that informs the underlying factor structure. Then, this inflation of correlations for the factors on the BOT-2 model indicates that the original model developed by the authors of the scale probably does not have four dimensions, since factors with high correlations should be combined for a parsimonious model, making it efficient with fewer resources. An exploratory structural equation modeling approach[27] could help identify the fit model for the BOT-2 (i.e., identify the number of independent factors).

For MABC-2, factorial validity for the original model developed by Henderson et al.[4] presented poor fit indices; despite being considered reasonable, in the present sample the multidimensional structure could not be confirmed for age band 1, in contrast with other studies.[6,8] The bifactor model, in turn, presented excellent fit indices, which indicates that the multidimensional model may not be the best one for MABC-2. Still, as shown by Hua et al.,[6] four tasks presented low factor loadings and needed to be excluded; therefore, there were less than three dimensions (subtests).

Regarding the results of the bifactor analysis for both the BOT-2 and MABC-2 models, we observed that the viability (ability to sustain the scores of the subscales
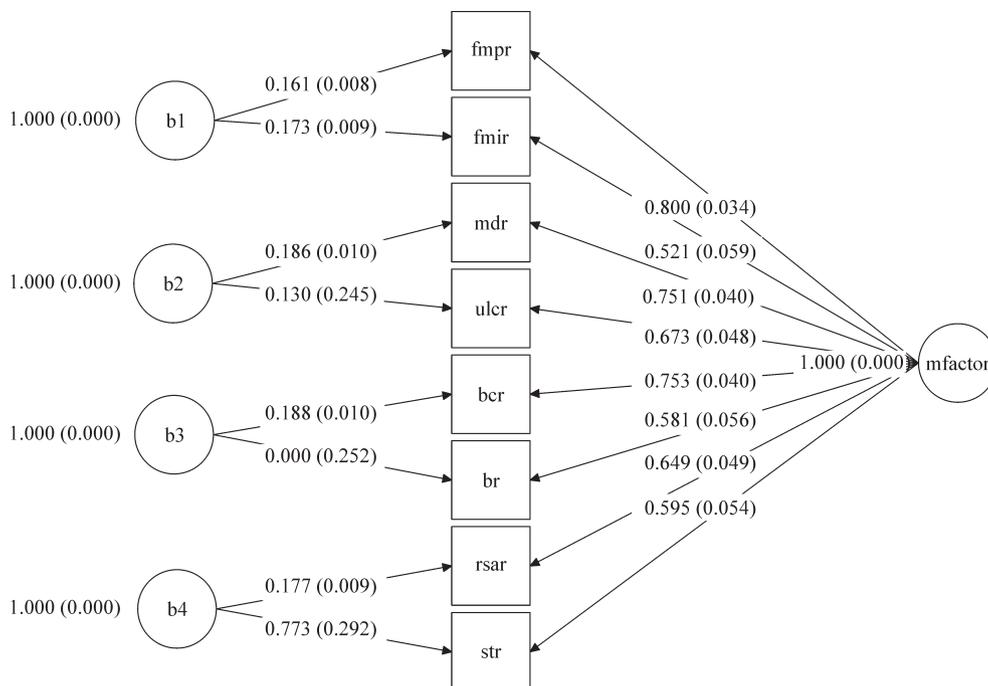
Figure 2 - Construct validity for the Bruininks-Oseretsky Test of Motor Proficiency, Second Edition (BOT-2) subscales. M-factor = general motor ability; b1 = fine manual control; b2 = manual coordination; b3 = body coordination; b4 = strength and agility; fmp = fine motor precision; fmi = fine motor integration; md = manual dexterity; ulc = upper-limb coordination; bc = bilateral coordination; b = balance; rsa = running speed and agility; st = strength; r = raw score.

after variation due to the M-factor) and reliability (the quality of the indicators) of the subscales were poor, and almost all the reliable variance in total scores could be attributed to the M-factor. This factor may reflect individual differences in motor performance. The M-factor is robustly reliable even though it is a multidimensional construct, and the specific subdomains displayed weak viability beyond the M-factor, indicating that the use of the subscales for BOT-2 and MABC-2 is not appropriate for estimation of motor skills; rather, the raw sum of item scores should be considered as an outcome measure. These results are consistent with the findings of the bifactor model applied to other areas of child evaluation and different scales assessing various aspects of psychopathology and personality.[20,24,28,29]

From a clinical perspective, these results suggest that the motor performance assessed by BOT-2 or MABC-2 reflects a general latent trait (M-factor); therefore, the reporting and interpretation of these tests should be restricted to the total motor composite for BOT-2 and the total score for the MABC-2. Moreover, the subtests should not be analyzed separately, since the values of the omegas of the subscales showed little reliable variation beyond the M-factor, limiting the use of the subtest scores as precise indicators of unique constructs.[14,20] Thus, in accordance with what has been suggested by the R-DoC,[3] when using these two instruments (BOT-2 and MABC-2) to verify the relationship between motor skills and pathophysiology in psychiatric and neurodevelopmental disorders, the most robust and precise way to verify and report motor performance is through the total score.

Limitations of the study

This research was conducted only with Brazilian students and did not address all the age groups proposed by BOT-2 and MABC-2, nor was the sample representative. Therefore, it is necessary to reproduce
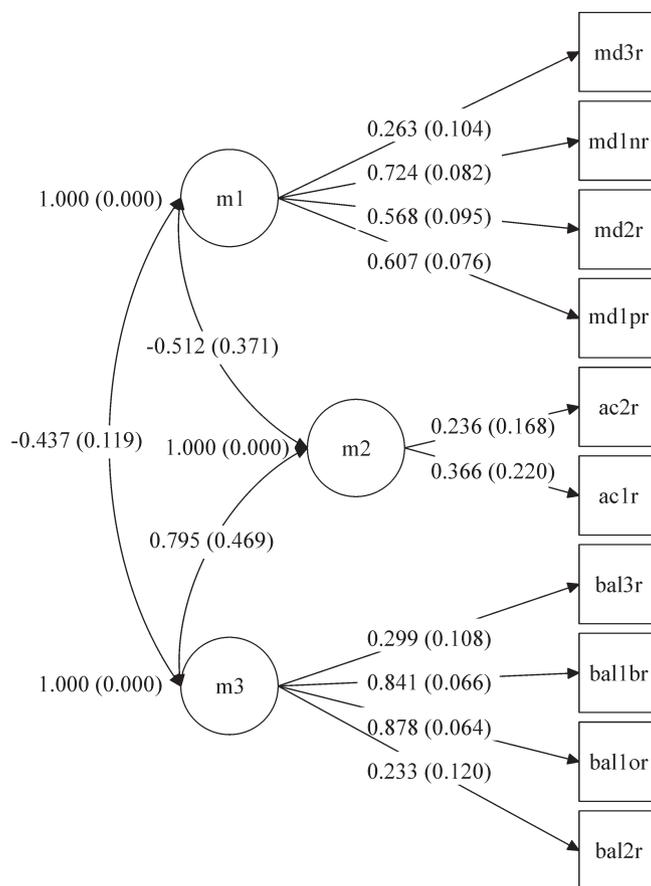
Figure 3 - Confirmatory factor analysis for the Movement Assessment Battery for Children, Second Edition (MABC-2). m1 = manual dexterity; m2 = aiming and catching; m3 = balance; md1p = posting coins with one's preferred hand; md1n = posting coins with one's non-preferred hand; md2 = threading beads; md3 = drawing trail; ac1 = catching beanbag; ac2 = throwing beanbag onto mat; bal1b = one-leg balance with one's "best" leg; bal1o = one-leg balance with one's other leg; bal2 = walking heels raised; bal3 = jumping on mats; r = raw score.

the analyses in other samples and age groups. However, the samples were robust, given the number of items in the tests evaluated, and the poor reliability and viability results found for the subscales are in accordance with recent and diverse studies using bifactor models across various areas of knowledge, including psychiatry and psychology.[20,24,28,29] Therefore, the results regarding the poor viability and reliability of the subscales are dependent of the sampling.

Conclusions

The BOT-2 exhibited factorial validity with a multidimensional structure in the current samples, but the MABC-2 presented poor fit indices, insufficient to confirm its multidimensional structure. Results of the bifactor model revealed that most of the reliable variance was derived from a general M-factor. Therefore, use of the BOT-2 and MABC-2 subscales is not supported in clinical or research settings. Even though our data are not representative of the entire country, this study is the first to use bifactor models with BOT-2 and MABC-2, and our findings address new insights regarding the use and interpretation of the assessment instruments most widely used with children.

Acknowledgments

Disclosure

No conflicts of interest declared concerning the publication of this article.
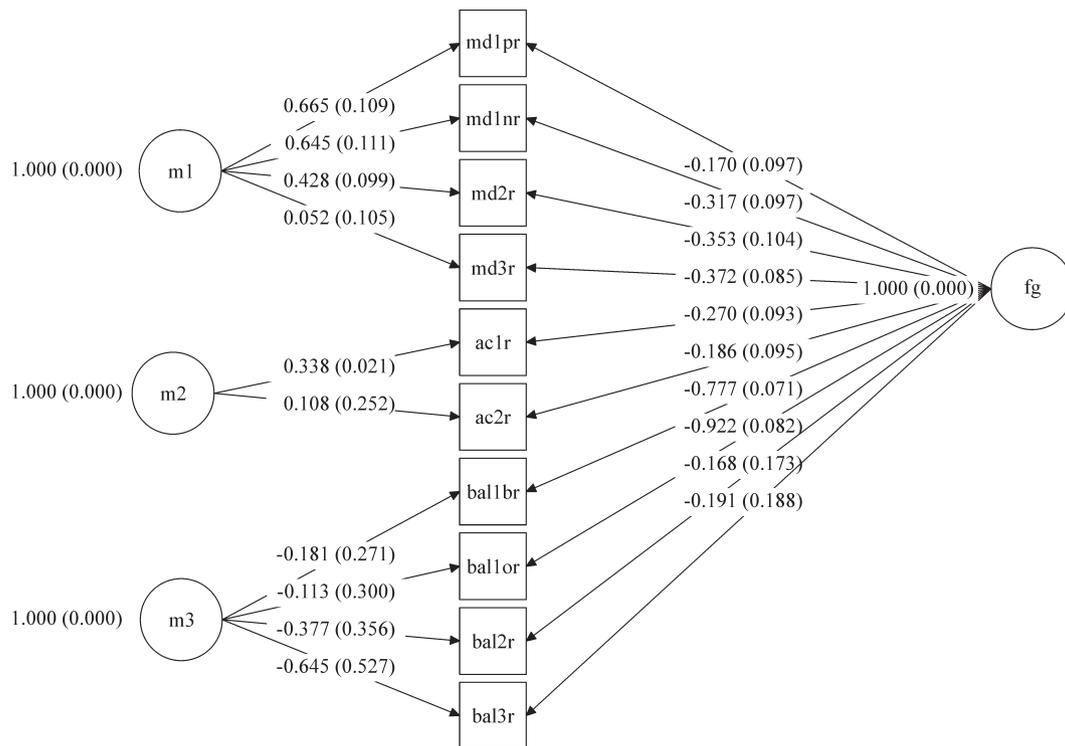
Figure 4 - Construct validity for the Movement Assessment Battery for Children, Second Edition (MABC-2) subscales. M-factor = general motor ability; m1 = manual dexterity; m2 = aiming and catching; m3 = balance; md1p = posting coins with one's preferred hand; md1n = posting coins with one's non-preferred hand; md2 = threading beads; md3 = drawing trail; ac1 = catching beanbag; ac2 = throwing beanbag onto mat; bal1b = one-leg balance with one's "best" leg; bal1o = one-leg balance with one's other leg; bal2 = walking heels raised; bal3 = jumping on mats; r = raw score.

References

1. American Psychiatric Association. Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition (DSM-5). Arlington: American Psychiatric Publishing; 2013.
2. Pasini A, D'agati E. Pathophysiology of NSS in ADHD. World J Biol Psychiatry. 2009;10:495-502.
3. Mittal VA, Bernard JA, Northoff G. What can different motor circuits tell us about psychosis? An RDoC perspective. Schizophr Bull. 2017;43:949-55.
4. Henderson SE, Sugden DA, Barnett AL. Movement Assessment Battery for Children-2: Movement ABC-2: Examiner's manual. London: Pearson; 2007.
5. Bruininks RH, Bruininks BD. BOT2: Bruininks-Oseretsky Test of Motor Proficiency: Manual. London: Pearson; 2005.
6. Hua J, Gu G, Meng W, Wu Z. Age band 1 of the Movement Assessment Battery for Children-Second Edition: exploring its usefulness in mainland China. Res Dev Disabil. 2013;34:801-8.
7. Schulz J, Henderson SE, Sugden DA, Barnett AL. Structural validity of the Movement ABC-2 test: factor structure comparisons across three age groups. Res Dev Disabil. 2011;32:1361-9.
8. Ellinoudis T, Evaggelinou C, Kourtessis T, Konstantinidou Z, Venetsanou F, Kambas A. Reliability and validity of age band 1 of the Movement Assessment Battery for Children--second edition. Res Dev Disabil. 2011;32:1046-51.
9. Kita Y, Suzuki K, Hirata S, Sakihara K, Inagaki M, Nakai A. Applicability of the Movement Assessment Battery for Children-Second Edition to Japanese children: a study of the age band 2. Brain Dev. 2016;38:706-13.
10. Valentini NC, Ramalho MH, Oliveira MA. Movement assessment battery for children-2: translation, reliability, and validity for Brazilian children. Res Dev Disabil. 2014;35:733-40.
11. Reise SP, Bonifay WE, Haviland MG. Scoring and modeling psychological measures in the presence of multidimensionality. J Pers Assess. 2013;95:129-40.
12. Stucky BD, Edelen MO. Using hierarchical IRT models to create unidimensional measures from multidimensional data. In: Reise SP, Revicki DA, editors. Multivariate applications series. Handbook of item response theory modeling: Applications to typical performance assessment. New York: Routledge/Taylor & Francis Group; 2015. p. 183-206.
13. Chen FF, West SG, Sousa KH. A comparison of bifactor and second-order models of quality of life. Multivariate Behav Res. 2006;41:189-225.
14. Reise SP. The rediscovery of bifactor measurement models. Multivariate Behav Res. 2012;47:667-96.
15. Yuan K-H, Bentler PM. On chi-square difference and z tests in mean and covariance structure analysis when the base model is misspecified. Educ Psychol Meas. 2004;64:737-57.
16. Nunnally JC, Bernstein IH, Berge JMT. Psychometric theory, New York: McGraw-Hill; 1967.
17. Brown TA. Confirmatory factor analysis for applied research. New York: Guilford Publications; 2015.
18. Muthén LK, Muthén BO. Mplus user's guide. 7th ed. Los Angeles: Muthén and Muthén; 2012.
19. Marsh HW, Hau K-T, Wen Z. In search of golden rules: comment on hypothesis-testing approaches to setting cutoff values for fit indexes and dangers in overgeneralizing Hu and Bentler's (1999) findings. Struct Equ Modeling. 2004;11:320-41.
20. Rodriguez A, Reise SP, Haviland MG. Evaluating bifactor models: calculating and interpreting statistical indices. Psychol Methods. 2016;21:137-50.
21. Reise SP. Invited paper: the rediscovery of bifactor measurement models. Multivariate Behav Res. 2012;47:667-96.
22. McDonald RP. Test theory: a unified treatment. Mahwah: Lawrence Erlbaum; 2013.
23. Revelle W, Zinbarg RE. Coefficients alpha, beta, omega, and the glb: comments on Sijtsma. Psychometrika. 2009;74:145-54.
24. Rodriguez A, Reise SP, Haviland MG. Applying bifactor statistical indices in the evaluation of psychological measures. J Pers Assess. 2016;98:223-37.
25. Reise SP, Horan WP, Blanchard JJ. The challenges of fitting an item response theory model to the Social Anhedonia Scale. J Pers Assess. 2011;93:213-24.
26. Raykov T, Marcoulides GA. An introduction to applied multivariate analysis. London: Routledge; 2012.
27. Asparouhov T, Muthén B. Exploratory structural equation modeling. Struct Equ Modeling. 2009;16:397-438.
28. Jovanović V. Structural validity of the Mental Health Continuum-Short Form: The bifactor model of emotional, social and psychological well-being. Pers Individ Dif. 2015;75:154-9.
29. Wagner F, Martel MM, Cogo-Moreira H, Maia CRM, Pan PM, Rohde LA, et al. Attention-deficit/hyperactivity disorder dimensionality: the reliable 'g'and the elusive's' dimensions. Eur Child Adolesc Psychiatry. 2016;25:83-90.

Correspondence:
Paola Matiko Martins Okuda
Rua Napoleão de Barros, 865 - Vila Clementino
04024-002 - São Paulo, SP - Brazil
Tel.: +55 (14) 981674959
E-mail: paolaokuda@yahoo.com.br