

# Development of the Brazilian version of the Child Hayling Test

## Desenvolvimento da versão brasileira do Teste Hayling Infantil

Larissa de Souza Siqueira,<sup>1</sup> Hosana Alves Gonçalves,<sup>1</sup>  
Lilian Cristine Hübner,<sup>2</sup> Rochele Paz Fonseca<sup>1</sup>

### Abstract

**Introduction:** The Hayling Test assesses the components of initiation, inhibition, cognitive flexibility and verbal speed by means of a sentence completion task. This study presents the process of developing the Brazilian version of the Child Hayling Test (CHT) and reports evidence of its content validity.

**Methods:** 139 people took part in the study. The adaptation was performed by seven translators and 12 specialist judges. An initial sample of 92 healthy children was recruited to test a selection of sentences adapted from previous adult and pediatric versions of the instrument, and a sample of 28 healthy children was recruited for pilot testing of the final version. The instrument was developed in seven stages: 1) translation, 2) back-translation, 3) comparison of translated versions, 4) preparation of new stimuli, 5) data collection with healthy children to analyze comprehension of the stimuli and analyses by the authors against the psycholinguistic criteria adopted, 6) analyses conducted by judges who are specialists in neuropsychology or linguistics, and 7) the pilot study.

**Results:** Twenty-four of the 72 sentences constructed were selected on the basis of 70-100% agreement between judges evaluating what they assessed and level of comprehensibility. The pilot study revealed better performance by older children, providing evidence of the instrument's sensitivity to developmental factors.

**Conclusions:** Future studies employing this version of CHT with clinical pediatric populations who have frontal lesions and dysfunctions and in related areas are needed to test functional and differential diagnoses of preserved or impaired executive functions.

**Keywords:** Neuropsychological tests, instrument development, Child Hayling Test, executive functions, inhibition, initiation.

### Resumo

**Introdução:** O Teste Hayling avalia os componentes iniciação, controle inibitório, flexibilidade cognitiva e velocidade verbal por meio de uma tarefa de completar frases. Este estudo apresenta o processo de desenvolvimento do Teste Hayling Infantil (THI) em português brasileiro, bem como evidências de validade de seu conteúdo.

**Métodos:** Participaram 139 sujeitos: A adaptação foi realizada por sete tradutores e 12 juízes especialistas. Uma amostra inicial de 92 crianças saudáveis foi recrutada para testar uma seleção de sentenças adaptadas de versões anteriores para adultos e crianças, e uma amostra de 28 crianças saudáveis foi recrutada para testar a versão final num estudo piloto. Sete etapas foram conduzidas: 1) tradução, 2) retrotradução, 3) comparação entre versões traduzidas, 4) desenvolvimento de novos estímulos, 5) coleta com crianças saudáveis para análise da compreensão dos estímulos e análise dos autores quanto aos critérios psicolinguísticos, 6) análise de juízes especialistas em neuropsicologia e linguística, e 7) estudo piloto.

**Resultados:** Das 72 frases construídas, selecionaram-se 24 que tiveram concordância de 70 a 100% entre juízes quanto aos construtos que avaliam e quanto ao grau de compreensibilidade. O estudo piloto revelou melhor desempenho nas crianças mais velhas, dando indícios da sensibilidade do instrumento aos fatores desenvolvimentais.

**Conclusões:** Estudos futuros que utilizem esta versão do THI em populações clínicas infantis que apresentem lesões e disfunções frontais e em áreas relacionadas tornam-se necessários para fins de diagnóstico funcional e diferencial de dificuldades executivas preservadas ou afetadas.

**Descritores:** Testes neuropsicológicos, desenvolvimento de instrumentos, funções executivas, inibição, iniciação.

### Introduction

The objective of this study is to describe the process of development of the Child Hayling Test (CHT) and present

evidence of its content validity. This version was inspired by the Brazilian and English versions for adults<sup>1,2</sup> and the Argentinean version of the Hayling Test, which is for Children.<sup>3</sup>

<sup>1</sup> Programa de Pós-Graduação em Psicologia, Grupo Neuropsicologia Clínica e Experimental, Pontifícia Universidade Católica do Rio Grande do Sul (PUCRS), Porto Alegre, RS, Brazil. <sup>2</sup> Programa de Pós-Graduação em Letras, Departamento de Estudos Linguísticos, PUCRS, Porto Alegre, RS, Brazil.

This paper is based on the first author's doctoral dissertation entitled "Teste Hayling Infantil: desenvolvimento, efeitos de idade e tipo de escola e evidências psicométricas," presented in 2014 at Pontifícia Universidade Católica do Rio Grande do Sul (PUCRS), Porto Alegre, RS, Brazil.

Financial support: Pró Bolsas, Pontifícia Universidade Católica do Rio Grande do Sul (PUCRS) – Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq). Submitted Feb 25 2015, accepted for publication Mar 09 2016. No conflicts of interest declared concerning the publication of this article.

Suggested citation: Siqueira LS, Gonçalves HA, Hübner LC, Fonseca RP. Development of the Brazilian version of the Child Hayling Test. Trends Psychiatry Psychother. 2016;38(3):164-174. <http://dx.doi.org/10.1590/2237-6089-2016-0019>

The Hayling Test is a two-part sentence completion task in which the last word of each sentence is missing. In part A, the task is to complete the sentence using a word that will result in the sentence making sense. In this case, both syntactic context and semantic context of the sentence are conducive to activation of a word that is frequently used and coherent in terms of the semantic field of the sentence. In part B, the word provided should not establish a meaningful relationship with the rest of the sentence. The examinee has to inhibit the dominant (logical, coherent) response and choose a word that is unrelated to both the syntactic context and the semantic context suggested by the incomplete sentence.<sup>1</sup> Versions have been adapted for administration to adults in several languages,<sup>1,2</sup> but, surprisingly, there appear to only be two versions for the pediatric population, an Argentinian<sup>3</sup> version and the Italian Junior Hayling Test.<sup>4</sup> There are also modified versions designed to be administered with neuroimaging.<sup>5,6</sup>

Considering that executive functions (EF) are a set of higher-order processes that serve to organize behavior and adapt it to context and to individual aims<sup>7,8</sup> and that these functions emerge in infancy and continue to develop through adolescence,<sup>9</sup> there is unquestionably a need for verbal tools to assess EF for clinical and research purposes.

This need is even more acute when it comes to instruments useful for investigating the EF components of initiation, inhibition and cognitive flexibility in a manner that is suited to the cultural, social and linguistic scenario in Brazil and particularly in samples of children. Executive function disorders can impact patients' functionality, affecting their social interactions and generating psychosocial and functional difficulties.<sup>8</sup> The Hayling Test (HT)<sup>1</sup> enables examination of the processes of initiation and inhibition in response to verbal stimuli. The complexity of these EF components and their relationships with oral language underscore the need for tools to be developed that are capable of examining controlled versus automatic processes in pediatric neuropsychological assessment.

The HT's specificity enables comparative assessment of several EF components. Using a single symbolic form (verbal), it is possible to detect dissociations in development of the components of initiation and inhibition with indexes of cognitive flexibility that can characterize neurological and psychiatric frames.<sup>10-12</sup> Studies have shown evidence of a double dissociation between specific EF disorders, such as in planning and response inhibition.<sup>13</sup> This highlights the clinical value of the HT. The Supervisory Attentional System (SAS) theory, developed by Norman & Shallice,<sup>14</sup> is the most acceptable theory for explaining the cognitive processes underlying accomplishment of parts A and B of the HT.<sup>15</sup> From the perspective of information processing, Norman & Shallice<sup>14</sup> adopted the distinction between automatic and controlled processing in parallel

to the cognitive processing demanded on a daily basis and in cases of problem solution in new situations.

The SAS is a set of cognitive processes involving initiation, goal-oriented strategy generation, and assessment of performance required to execute complex cognitive tasks not used on a daily basis or nor learned yet. This system is activated when automatic actuation of familiar cues is unable to select frames of thoughts and of automatic actions that can effectively accomplish a target. In this case, automatic actions of sequential thoughts are inadequate or insufficient, demanding new behaviors to be planned and monitored to achieve satisfactory performance in the task.<sup>14,16</sup>

Within a cognitive neuropsychological perspective, it is acknowledged that the HT provides a measure of the capacity to create strategies targeting the accomplishment of tasks.<sup>3,17,18</sup> Incapacity to generate adequate strategies causes poor performance in the HT in the form of errors.<sup>18</sup> Moreover, administration of neuropsychological tests to assess EF is important for analysis of the components involved at different developmental ages.

Studies with samples of children with typical development that have administered instruments to measure EF, including the HT, have shown significant age effects on all subitems of executive performance, particularly on inhibition.<sup>19</sup> In case studies of autistic spectrum disorders, researchers have observed impairment in planning, inhibition and self-monitoring of EF tasks.<sup>20</sup> Significantly lower performance on the HT has been reported in studies with children who have difficulties with text comprehension, suggesting that the relationship between reading comprehension and working memory could rely on abilities to inhibit irrelevant information.<sup>21,22</sup>

It is not always easy to consistently identify deficits indicative of abnormal executive processes using validated tests representing daily behavior.<sup>23</sup> Studies with children with attention deficit hyperactivity disorder (ADHD) point to abnormalities in the domain of inhibitory and attentional EF processes, such as generation and use of strategies and behavioral control.<sup>4,24-26</sup>

In addition to the HT's sensitivity for assessment of children with ADHD, it is also an important instrument for assessing executive impairment in schizophrenia,<sup>27</sup> in autism,<sup>20</sup> and within other clinical frames,<sup>28,29</sup> demonstrating its usefulness for diagnosis of executive and cognitive dysfunctions. However, research with children with typical development is still needed, since such studies are crucial to defining norms and to identifying biological and sociocultural factors related to cognitive and executive processes.

Studies with children in late childhood using the adult version of the HT<sup>2</sup> and other EF assessment tasks suggest the existence of correlations between measures of EF

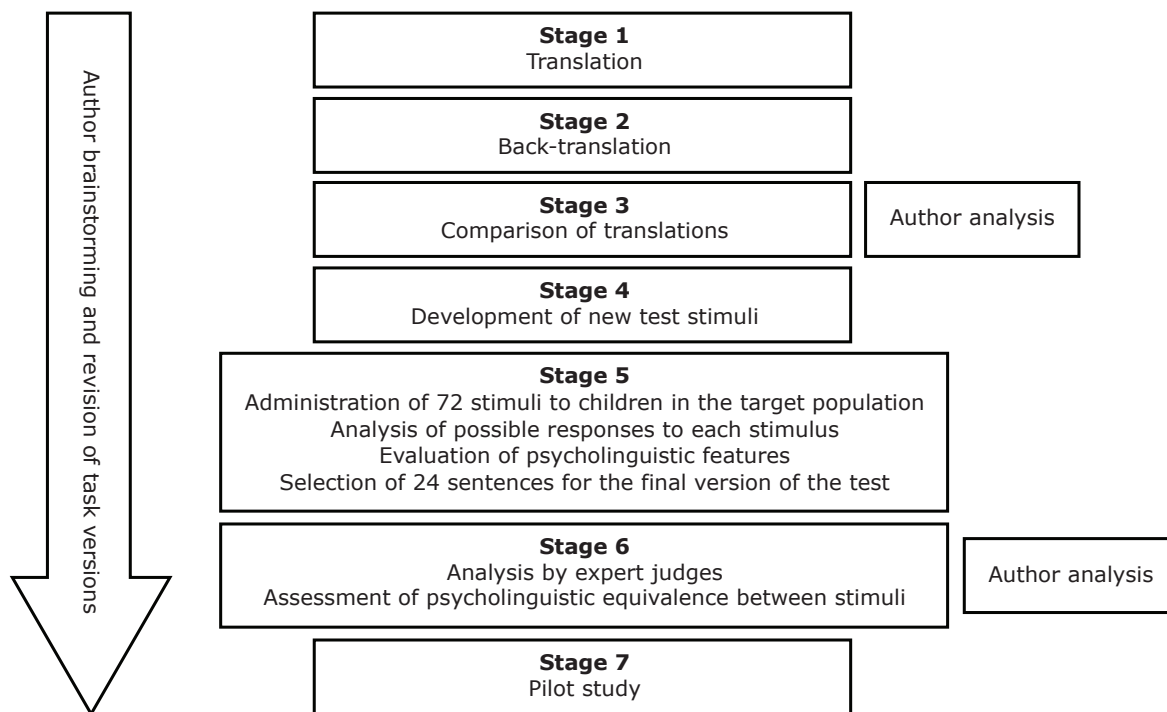
components, particularly between scores for inhibition and cognitive flexibility.<sup>30</sup> However, another study<sup>31</sup> has recommended adaptation of the test for samples of children, since the adult version proved to have little sensitivity for differentiating groups of school-aged participants.

In common with the HT, there are other nationally and internationally known instruments for neuropsychological assessment that are intended for investigation of executive processes such as initiation, inhibition, processing speed and cognitive flexibility. Among these, the most widely adopted are founded on verbal paradigms, measures of verbal fluency, go no go type tests, digit span, and n-back verbal-auditory.<sup>32,33</sup> However, due to the fact that none of them are two-way like the HT, identification of dissociations between automatic and controlled processes is a challenge, which is overcome by the HT. The Trail Making Test (TMT) is an interesting equivalent to the HT that is based on a non-verbal paradigm.<sup>34</sup>

As can be observed from the review presented here, making available verbal instruments for neuropsychological evaluation that are quick and easy to administer, as is the case of the HT, should contribute to refinement of the profile of cognitive functions, more specifically, of the subprocesses of EF. Moreover, adaptation of this instrument for Brazilian Portuguese-speaking school-age samples should be of help for development of more accurate diagnoses of executive dysfunctions, for prognostics, for planning and execution of training programs targeted at teachers and clinicians, and for neurocognitive rehabilitation.

## Methods

This study was conducted in 7 stages, each of which employed specific samples and procedures. Figure 1 illustrates the stages of development of the CHT.



**Figure 1** - Flowchart illustrating development of the CHT.

## Sample

The entire sample comprised one hundred and thirty-nine participants, as follows. Seven translators took part in the general translation procedures, 12 specialists acted as judges in the assessment stage,

92 healthy children representing the target population took part in a data collection stage, and 28 healthy children were recruited for the pilot study.

Two translators participated in stage 1, forward translation. Two bilingual translators who were native speakers of Brazilian Portuguese and had high-level

proficiency in Spanish and/or English were involved in stage 2, back-translation. Three more translators participated in stage 3, comparison of translations. In stage 4, the authors of the instrument constructed new stimuli, using books for children and school primers to ensure that vocabulary appropriate for samples of children was used.

In stage 5, 92 healthy children participated, 56 of whom were studying at public schools and 36 of whom were at private schools. Their mean age was 8.90 (standard deviation [SD] = 1.87) and they had a mean Raven percentile of 83.23 (SD = 15.66) for intellectual abilities. These participants exhibited low rates of hyperactivity and/or inattention, according to Conners' Abbreviated Questionnaire scores (mean [M] = 3.45; SD = 4.72).

Eight judges specialized in neuropsychology and four specialized in psycholinguistics participated in stage 6, analysis by specialized judges. In the pilot study (stage 7), 28 children with typical development were assessed. They were evenly distributed according to age, school year and type of school, as follows: two children for each year of age (from 6 to 12 years old) and for each school grade (from the 1st to the 7th) and one child from each type of school (public and private), making a total of 28 healthy children in the sample. The only exception to this distribution was that four 6-year-old children were included because whereas children of other ages could be enrolled in different school grades (for instance, seven-year-old children could be in the 1st or the 2nd grades), six-year-old children could not be enrolled in any other grade than the first. This made it possible to meet the inclusion criterion of being in at least the 1st school year.

### Procedures and instruments

Firstly, the authors of the original instrument, the initial adult version, were contacted.<sup>35</sup> They gave permission for development of this Brazilian version for children. Although there are two child versions – the Italian and the Argentinean – access was only obtained to the stimuli from the second of these. The seven stages conducted to develop the adapted instrument, as illustrated in Figure 1, are described below.

**Translation; back-translation; comparison of translations; and linguistic adjustments to the instructions, scoring norms and stimuli (stages 1, 2 and 3).** The versions for adults (English)<sup>1</sup> and for children (Argentinean)<sup>3</sup> underwent the processes of translation, back-translation, comparison of

translations and linguistic adjustments of the test instructions. Working from the Brazilian Portuguese adult version<sup>2</sup> as a starting point, adjustments were made to the instructions for the two parts (A and B) and to the scoring norms to adapt them for the pediatric population.

The two international versions were each translated by a translator with attested proficiency in the languages involved. The back-translations were conducted by two bilinguals who were native speakers of Brazilian Portuguese and had advanced levels of proficiency in the foreign languages (one in Spanish, the other in English). The authors of the Junior Hayling Test<sup>4</sup> were contacted and gave their permission for their version to be used as inspiration for development of the Brazilian child version, but only the adult version in English<sup>1</sup> and the child version in Spanish<sup>3</sup> were translated.

Three translators participated in the comparison of translations. Two of them acted as independent judges of the initial Brazilian Portuguese versions and the third translator was used to arrive at a consensus on stimuli and instructions on which the two judges had not reached a consensus.

**Construction of new stimuli (stage 4).** The authors prepared new stimuli (sentences) and implemented syntactic, semantic and pragmatic adjustments to the instructions for the first version of the CHT, in accordance with the original Brazilian and English adult versions<sup>1,2</sup> and/or the Argentinean version.<sup>3</sup> This preparation process adhered to certain criteria, including syntactic and semantic judgments on the sentences, and adaptation to fit the stages of syntactic development in childhood, the vocabulary frequency of Brazilian children, and the morphosyntactic features of Brazilian Portuguese. To achieve this, alternative stimuli were sought in materials such as booklets and books for children. In contrast with the adult version,<sup>2</sup> in which nouns, adjectives, and adverbs are accepted to coherently complete the sentences, the stimuli chosen for the children's version are completed exclusively using nouns. The aim of this stage was to meet linguistic criteria for adaptation to the sociocultural environment and to Brazilian children's developmental phases. At the end of this stage, 72 sentences had been prepared for a pilot version. Of these, four were based on the Argentinean version for children and two on the adult Brazilian version, in terms of semantic and/or syntactic aspects (Table 1). The authors of the instrument, who are psycholinguists, psychologists, neuropsychologists and speech pathologists, created the remaining sentences.

**Table 1** - Sentences adapted from previous versions of the Hayling Test

Original sentence	Corresponding sentence in the Child Hayling Test*
Argentine version <sup>3</sup>	
Julián took an aspirin for the pain in his...	The grandfather took some medicine for the pain in his...
As soon as the fire started, they called the...	The firefighter put out the...
The boys were reading a new...	The student chose a story book from the...
Before going to sleep, the woman turned off all the...	Before going to sleep, the child put on a...
Brazilian version <sup>2</sup>	
The football player made a...	The football player made a...
The boy kicked the ball at the...	
The cook put the cake in the...	The lady put the cookies in a...

\* These examples have all been translated into English "literally."

**Administration of the 72 stimuli to healthy children representing the target population; identification of possible answers and judgment of the psycholinguistic criteria of the sentences; selection of 24 sentences for the final version; and implementation of a study for the analysis of new stimuli (stage 5).** A sample of 92 children of both genders was selected to participate in this phase of instrument development. Their ages ranged from 6 to 12 and they were studying at public or private schools and therefore corresponded to the profile of the target population for the Brazilian version of the CHT. These children's parents signed free and informed consent forms and filled out a questionnaire providing sociodemographic and health-related information.<sup>36</sup> This instrument was useful for selection of participants, since it allowed researchers to identify children with neurological and/or psychiatric disorders, or with uncorrected sensory problems, among other exclusion criteria. Teachers answered Conners' Abbreviated Questionnaire,<sup>37</sup> with the objective of identifying children with signs of possible inattention and hyperactivity in the classroom.

The first set of sentences prepared as part of development of the CHT were administered to the 92 children recruited. The 6 to 7-year-old children were asked to say two words that made sense with the content of each sentence, while the older children were asked to answer in writing. Sentences were administered to the 6 to 7-year-old children individually and to the 8 to 12-year-old children in groups. Raven's Progressive Matrices Test<sup>38</sup> was also administered to these participants in order to exclude children with evidence of possible intellectual deficits. After this stage, the frequencies of answers to each sentence were analyzed using the Statistical Package for the Social Sciences (SPSS), version 17.0, and the sentences were classified. Psycholinguistic criteria were observed, such as the level of semantic constraint in terms of sentence structure (for instance, agreement in number and gender), as well as the meanings of

the words provided as answers. Additionally, in order to guarantee that the instrument achieved its aim (examining verbal executive components), the syntactic complexity and extension, whether verbs were transitive or intransitive and whether they took direct and/or indirect objects were also considered. Twenty-four of the 72 sentences prepared for the pilot study were selected for the final version of the instrument. The varying levels of syntactic complexity of the sentences from the original instruments were maintained. Attention was also devoted to balancing syntactic structures between parts A and B, to control the linguistic and cognitive demands involved in processing of sentence completion in each of the parts.

**Analysis by specialized judges (stage 6).**

Psychologists, speech pathologists and linguists specialized in neuropsychology analyzed the sets of instructions and stimuli. In the first stage, the judges gave an opinion on which neuropsychological or EF components were predominantly being assessed. The judges were instructed to number one or more cognitive components and state which components were predominantly being assessed by the sets of instruction and stimuli for parts A and B. The components they identified were initiation, inhibition, processing speed, cognitive flexibility and lexical-semantic processing. In the second stage, the specialist judges assessed the stimuli's adequacy for the tasks and whether they represented the processes assessed. For this stage, 11 pairs of sentences were judged, one of which was eliminated after the judges gave their opinions. At the end, the judges answered two questions asking whether the sentences would be easily understood.

Those stimuli with at least 80% agreement from the judges were included. Therefore, the instrument was evaluated according to the executive components it assessed, to its adequacy for the target population, to the comprehensibility of the stimuli, and to its representativeness of the processes assessed. It was also

evaluated according to the lengths in number of words of its stimuli and in terms of the syntactic equivalence between pairs of sentences, to compare the two parts of the test. Additionally, the judges made suggestions for improvements to the test instructions.

**Pilot study with administration of the final version of the instrument to 28 children aged 6 to 12, from public and private schools (stage 7).** Twenty-four of the sentences from the pilot study were selected for the final version. After analysis of the judges' considerations and suggestions by the authors, the final version of the test was individually administered to a pilot sample composed of 28 children (characterized in the "sample" subsection above). The objective of the pilot study was to estimate the time required for administration and verify comprehension of the instructions and the children's answers. Moreover, the pilot study provided the authors with preliminary evidence on the level of difficulty of the instrument as a whole, helping to avoid ceiling and floor effects.<sup>39</sup> Modifications were not necessary, since the 28 participants did not exhibit difficulties with completing the test. At this point, development of the CHT version to be normalized was concluded.

### **Data analysis**

In the step shown in stage 5 of the flow chart, SPSS 17.0 was used to conduct an analysis of the frequency of answers provided for each of the sentences. The sentences were then classified according to these answers on the basis of semantic restriction, syntactic complexity and sentence length. Based on the distribution of the frequencies of occurrence of the various words and semantic fields in each sentence, the authors empirically classified each sentence as "restrictive" (those which strongly restricted the word most often chosen to complete the sentence), "intermediately restrictive" or "weakly restrictive" (sentences with a higher number of possible answers).

A descriptive statistical analysis was conducted with percentages of answers and calculation of the rate of agreement between judges. The number of items on which the judges agreed was divided by the total number of items, considering the minimal cutoff of 80% agreement, to provide evidence of content validity.<sup>40,41</sup>

### **Results**

The results are presented according to the stages of the process to develop the CHT.

### **Stages 1, 2 and 3: translation, back-translation, and comparison of translations**

The instructions for the CHT versions generated by translation and back-translation were adapted for Brazilian Portuguese. These translated versions exhibited similarities, such as 15 sentences in each of the parts (A and B), sentences with varying lengths, ending with articles, prepositions, adjectives, pronouns, or verbs.

### **Stage 4: construction of new stimuli**

The authors of the CHT created new stimuli based on the criteria of familiarity and ease of sentence completion. The sentences were inspired by the stimuli in the existing instruments for children (the Junior Hayling Test; the Spanish version of the Hayling Test for children)<sup>3,4</sup> and by day-to-day topics that might interest children aged 6 to 12. The stimuli should also offer the sensitivity needed to diagnose executive impairment. The instructions based on the existing instruments were maintained, with some adaptations to suit the Brazilian children's sociolinguistic situation.

### **Stage 5: administration of the 72 stimuli to healthy children representing the target population**

Each sentence was classified based on a restriction level accordingly to the frequency of answers given to each one of them. Table 2 summarizes how the 72 sentences from the initial pilot study were classified.

After this analysis, 24 sentences were selected for the final version and classified in order to establish sentence equivalence between parts A and B. Criteria of semantic restriction, and syntactic extension and complexity were adopted, and the criteria used for the Junior Hayling Test<sup>4</sup> were respected.

Part A comprises 10 incomplete sentences that the child has to complete with a word that makes sense within the sentence context, plus two sentences for practice. Part B comprises 10 incomplete sentences to be completed by the child with a word that does not make sense within the sentence context and, in common with Part A, it also contains 2 practice sentences.

### **Stage 6: specialized judges' analyses**

The rates of agreement between the specialized judge for the two stages are shown in Tables 3 and 4.

The majority of the sentences proposed for Part A were judged by the specialized judges to be assessing initiation. In a

similar manner, the specialized judges evaluated the majority of the sentences proposed for Part B as assessing inhibition.

However, sentence 8 for Part B only achieved an agreement index of 60.0%, resulting in exclusion of this sentence.

**Table 2** - Restriction level of the sentences of the CHT pilot study

Level of sentence constraint	Definition	Percentage of sentences	Sample sentence
High constraint 1	All responses within a single semantic field, most consisting of the same word, which accounts for at least 80% of responses.	22.53	Before going to sleep, the child put on a... On her wedding day, the bride wears a...
High constraint 2	All responses within a single semantic field, consisting of words which are produced with a more uniform frequency distribution.	11.16	The grandfather took some medicine for the pain on his... The football player kicked the ball at the...
Intermediate constraint 1	All responses within one of two semantic fields, one of which accounts for at least 80% of responses.	35.21	The farmer milks the... The boy poured the juice in a...
Intermediate constraint 2	All responses within one of two semantic fields, which account for a roughly similar proportion of responses (50%).	12.68	The lady put the cookies in a... The man keeps his money in a...
Low constraint 1	Most responses within one semantic field (50%), with the remaining responses arising from several other categories.	12.68	The boy bit the... The storm knocked over the...
Low constraint 2	Responses arising from several semantic fields, none of which predominates over others.	5.74	In my room, there are many... On my birthday, there were many...

**Table 3** - Result of specialized judges' analyses of the cognitive/executive function assessed by each stimuli

Stimuli	Cognitive/ executive function evaluated	Agreement (%)
Part A – sentences 1 to 5	Initiation	90.0
Part A – sentences 6, 7 and 11	Initiation	80.0
Part A – sentences 8 to 10	Initiation	70.0
Part B – sentences 1 to 4	Inhibition	80.0
Part B – sentences 5 to 7, 9 and 11	Inhibition	70.0
Part B – sentence 8	Inhibition	60.0

**Table 4** - Result of specialized judges' analysis concerning the comprehensibility of instructions and stimuli

Stimuli	Agreement (%)
Part A – sentences 1, 2, 4, 6 to 10	100.0
Part A – sentences 3, 5 and 11	90.0
Part B – sentences 1 and 2	90.0
Part B – sentences 3 to 11	100.0
Sentences adequate	90.0
Instructions adequate	90.0

Based on the specialized judges' analyses and on their suggestions, stimuli and instructions were revised and modifications were made. All changes were intended to make the sentences administrable to children aged 6 to 12. Sentences in Part A underwent syntactic reduction, deletion of adjectives, and changes to articles and verbs, and a similar process was applied to Part B. However, care was taken to maintain the criteria of control over syntactic and lexical-semantic structures and equivalence between Part A and B as well as in relation to the original test.

The changes made to the CHT's instructions were intended to make the test easier to understand, in addition to emphasizing the tasks' rules (completing sentences quickly, saying only one word, and the differences between Parts A and B).

### Stage 7: pilot study

Table 5 lists descriptive data for the results of administration of the final version of the CHT in the pilot study, showing scores indicating the sample's performance.

The data presented in Table 5 show that participants had no difficulties with performing the test. In general, the results of the pilot study demonstrate that the instructions were well understood by the participants and that their answers seem to have been as expected for their ages and the aims of the instrument. The average time taken for administration and scoring was 10 minutes. Considering performance according to

age, the participants in the pilot study exhibited similar performance in terms of time and errors in both parts of the CHT. However, in Part B there is preliminary evidence of gradual improvement as age increases, mainly in the time variable. Regarding type of school, children from private schools made fewer mistakes and accomplished the task in a shorter time than those from public schools.

**Table 5** - Participants' performance in the pilot study of the Child Hayling Test, broken down by age and type of school (n = 28)

	Variables in the Child Hayling Test (mean ± standard deviation)			
	Part A - time	Part A - errors	Part B - time	Part B - errors
Age				
6 years (n = 4)	24.66±6.83	0.75±0.96	44.25±5.34	6.00±2.58
7 years (n = 4)	17.45±4.80	0±0	40.70±7.12	4.50±0.70
8 years (n = 4)	20.68±8.01	0±0	42.27±11.69	7.50±0.70
9 years (n = 4)	14.13±5.37	0±0	56.92±55.05	3.00±1.41
10 years (n = 4)	11.07±0.95	0±0	37.49±4.99	5.00±2.83
11 years (n = 4)	11.07±0.95	0±0	37.49±4.99	2.00±0
12 years (n = 4)	15.38±0.80	0±0	27.77±3.69	1.50±0.70
Type of school				
Private (n = 14)	16.99±4.83	0±0	41.67±23.66	4.13±2.16
Public (n = 14)	18.52±8.07	0.50±0.76	38.69±10.94	4.75±2.92

## Discussion

This article describes the process of developing a standardized task, designed according to rigorous criteria to assess the EF components of initiation and inhibition. The results indicate that this process met psychometric criteria that ensure the instrument's validity.

There is a growing concern among clinicians and researchers with developing and validating refined tests to assess performance in specific cognitive tasks.<sup>39,42</sup> More specifically, attention has been dedicated to stimuli selection, to modification and to adaptation for specific populations, taking their cultural, social and socioeconomic status into consideration.<sup>43</sup>

The project to develop the CHT is part of the field of research devoted to investment in new resources for diagnosis of cognitive abilities. This rich and complex process allows professionals to arrive at neuropsychological construct assessments, besides enabling information exchange among scientific communities, including in international studies.<sup>44</sup>

Despite growing concern with administration of neuropsychological tasks that are more adequate to each population sample, as far as we know there is no standardization of the stages that should be performed in a cross-cultural adaptation or in construction of

new instruments. This is a hindrance to choosing the methodological procedures to be conducted and is a difficulty that was mentioned by Argollo et al.<sup>44</sup> when describing the process of cross-cultural adaptation of a neuropsychological battery for children - NEPSY (A Developmental NEuroPSYchological Assessment). As a result, the stages in development of the children's version of the HT were based on criteria adopted by the authors, taking linguistic and cultural aspects into consideration, as well as developmental features of Brazilian children.

A systematic review conducted by Fonseca et al.<sup>39</sup> analyzed the process of adapting verbal instruments and found that the stages commonly adopted are translation, back-translation, judgments by specialists (professional experts in the fields encompassed by the instrument) and administration in pilot studies. However, some studies exhibit greater concern with ensuring the adequacy of stimuli to the sociolinguistic and cultural conditions of the country in which the adapted instrument will be administered. In the process to develop the CHT, these aspects were considered extremely necessary, since the instrument was adapted from international versions. In addition to attending to these criteria, care was also taken with psycholinguistic criteria to guarantee equivalence of stimuli.



Methodological rigor is essential when it comes to the linguistic aspects of instrument adaptation, i.e., when linguistic stimuli are used to examine cognitive functions. This is the case because linguistic and cultural specificities can limit the administration and further analysis of some tests.<sup>45,46</sup> The CHT appears to have been adapted well, according to the psycholinguistic criteria described by Altmann<sup>47</sup> and Berman.<sup>48</sup>

In the process of developing the CHT, the translation procedures employed included direct and compared translation.<sup>49</sup> Although translation is in general the first procedure conducted in an instrument adaptation process, it is not alone sufficient to guarantee the adapted version's adequacy. Acevedo et al.<sup>50</sup> proposed general guidance for the process of adaptation of verbal and non-verbal neuropsychological instruments to Spanish, emphasizing the importance of distinguishing literal translation and adaptation. This is the case because direct translation results in non-adjusted expressions, which may harm the measurement of the target cognitive components. Similarly, it is very important to adapt the instructions as well, i.e., in this case they must be modified in order to make them more comprehensible to children, without altering the construct assessed.<sup>51</sup>

According to Wilson,<sup>52</sup> neuropsychological batteries do not favor the control or manipulation of variables such as the stimuli used for assessment, including, for instance, lexical frequency, familiarity, and grammatical category. Janczura et al.<sup>53</sup> and Stein & Gomes<sup>54</sup> also emphasize the importance of standardization using familiarity lists for Brazilian Portuguese. These authors mention that rules of concreteness, words' semantic associations, and word frequencies, among others, allow the researcher to advance cognitive function studies, using adequate materials and, consequently, refining the quality and reliability of results. In this context, the stage in which 72 candidate stimuli were administered to healthy children representing the target population was considered necessary to collect relevant psycholinguistic criteria for each sentence in the instrument. By means of this stage it was possible to classify the sentences created for the CHT according to the level of restriction of answers and, posteriorly, detect which sentences were appropriate for subsequent stages in the development and adaptation process.

The stage of analysis by specialist judges was included because it is a procedure commonly found in adaptation studies, as in the cross-cultural adaptations of the NEPSY Battery,<sup>44</sup> the Wechsler Adult Intelligence Scale, third edition (WAIS-III) and Wechsler Intelligence Scale for Children, third edition (WISC-III) Scales,<sup>55</sup> and the Brief Battery of Communication Assessment MAC.<sup>51</sup> The analysis by specialist judges can be compared to

the theoretical analysis of items that, in psychometry, according to Pasquali,<sup>56</sup> aims to evaluate the hypothesis that each item adequately represents the construct.

To supplement the test development process, the pilot study was conducted to verify whether it was necessary to alter sentences and to provide an estimated total administration time. Discriminative and quick to administrate tasks are clinically useful for assessment of school age children. According to Fonseca et al.,<sup>39</sup> the literature does not indicate the ideal sample size for a pilot study. This CHT study was intended to assess healthy children with the ages and educational characteristics of those who would take part in future normative samples.

In this context, the results of this instrument development process show which words and/or expressions had to be modified to be made suitable for the children's linguistic knowledge, leading to better sentence comprehension. Moreover, the analyses by a neuropsycholinguist and by specialized judges were of fundamental importance. These procedures ensured that the aims of the original instruments were maintained, while serving to assess whether the sentences were appropriate for the samples of children. Finally, the pilot study completed the process, providing empirical data from a real assessment context, testing the task with children in the target age groups. This stage contributed evidence about the applicability of the task and led to final adjustments, as was the case in a study by Bandeira et al.<sup>57</sup> While demonstrating the roles played by age and type of school in the EF components assessed by the CHT, these data showed that children's performance improved as age increased. This fact may be related to improved attentional, executive and linguistic capacities.<sup>44,58,59</sup> The preliminary results of cases of each age indicate that the task seems to adequately differentiate performance levels, with a potential to aid in identification of children with typical and atypical development.

Regarding school type, the pilot study suggests children from private schools perform better on the CHT stimuli. This result corroborates the findings of studies that claim there is an influence from the sociocultural conditions on the level of cognitive stimulation offered to children enrolled at this type of school, possibly both at school and at home. Therefore, education plays an important role in children's development and is related to EF improvement.<sup>60,61</sup>

One limitation of this study is that it is not possible to determine whether the number of participants in the pilot study was sufficient, since there is no consensus in the literature regarding an adequate sample size for adaptation and development of neuropsychological instruments. It is suggested that a systematic literature

review be conducted of pilot studies presenting sensitivity data and analyzing sample size in adaptation of instruments for children.

Despite this limitation, the CHT offers implications for the functional diagnosis of executive components. Its two-part format enables associations and dissociations in performance to be identified. The CHT adaptation process was successful because it comprised seven complex and rigorous stages. Each of these steps was useful to maintain the HT's score interpretations, offering evidence of the test's content validity. Future studies will focus on psychometric parameters to define reliability, criterion validity and construct validity and on acquiring normative data by age and school type. Studies of the relationships between performance in the HT and in other neuropsychological tasks will also be needed, attempting to identify diagnostic associations and dissociations, in the light of cognitive neuropsychology. Similarly, clinical studies of sensitivity and of specificity should be conducted with neurological and psychiatric infant populations.

## References

- Burgess PW, Shallice T. The Hayling and Brixton Tests. Bury St. Edmunds: Thames Valley Test; 1997.
- Fonseca RP, Oliveira C, Gindri G, Zimmermann N, Reppold C. Teste Hayling: um instrumento de avaliação de componentes das funções executivas. In: Hutz CS. Avaliação neuropsicológica de crianças e adolescentes. São Paulo: Casa do Psicólogo; 2010. p. 337-64.
- Abusamra V, Miranda MA, Ferreres A. Evaluación de la iniciación e inhibición verbal en español: adaptación y normas del test de Hayling. *Rev Arg Neuropsicol.* 2007;9:19-32.
- Shallice T, Marzocchi GM, Coser S, Del Savio M, Meuter RF, Rumiat RI. Executive function profile of children with attention deficit hyperactivity disorder. *Dev Neuropsychol.* 2002;21:43-71.
- Collette F, Van der Linden M, Delfiore G, Degueldre C, Luxen A, Salmon E. The functional anatomy of inhibition processes investigated with the Hayling task. *Neuroimage.* 2001;14:258-67.
- Robinson GA, Cipolotti L, Walker DG, Biggs V, Bozzali M, Shallice T. Verbal suppression and strategy use: a role for the right lateral prefrontal cortex? *Brain.* 2015;138:1084-96.
- Bielak AA, Mansueti L, Strauss E, Dixon RA. Performance on the Hayling and Brixton tests in older adults: norms and correlates. *Arch Clin Neuropsychol.* 2006;21:141-9.
- Diamond A. Executive functions. *Annu Rev Psychol.* 2013;64:135-68.
- Sun J, Mohay H, O'Callaghan M. A comparison of executive function in very preterm and term infants at 8 months corrected age. *Early Hum Dev.* 2009;85:225-30.
- Kraan CM, Hocking DR, Georgiou-Karistianis N, Metcalfe SA, Archibald AD, Fielding J, et al. Impaired response inhibition is associated with self-reported symptoms of depression, anxiety, and ADHD in female FMR1 premutation carriers. *Am J Med Genet B Neuropsychiatr Genet.* 2014;165B:41-51.
- Robinson G, Walker D, Cipolotti L, Biggs V, Bozzali M, Shallice T. Prefrontal contributions to initiation, suppression and strategy: a neuropsychological study of focal frontal patients. *Front Hum Neurosci.* 2015;9.
- Whitton AE, Henry JD, Grisham JR. Moral rigidity in obsessive-compulsive disorder: do abnormalities in inhibitory control, cognitive flexibility and disgust play a role? *J Behav Ther Exp Psychiatry.* 2014;45:152-9.
- Johnston K, Madden AK, Bramham J, Russell AJ. Response inhibition in adults with autism spectrum disorder compared to attention deficit/hyperactivity disorder. *J Autism Dev Disord.* 2011;41:903-12.
- Norman DA, Shallice T. Attention to action: willed and automatic control of behavior. In: Davidson RJ, Schwartz GE, Shapiro D, editors. *Advances in research and theory.* New York: Plenum; 1986. p. 1-18.
- Chan RC, Shum D, Touloupoulou T, Chen EY. Assessment of executive functions: review of instruments and identification of critical issues. *Arch Clin Neuropsychol.* 2008;23:201-16.
- Chan KK, Xu JQ, Liu KC, Hui CL, Wong GH, Chen EY. Executive function in first-episode schizophrenia: a three-year prospective study of the Hayling Sentence Completion Test. *Schizophr Res.* 2012;135:62-7.
- Cartoceti RV, Sampedro B, Abusamra V, Ferreres A. Evaluación de la iniciación y la supresión de respuesta verbal en niños. Versión infantil en español del Test de Hayling. *Rev Fonoaudiol.* 2009;55:9-24.
- Abusamra V, Miranda MA, Ferreres A. Test para evaluar la iniciación e inhibición verbal. Adaptación al español del test de hayling. In: XIII Jornadas de Investigación y Segundo Encuentro de Investigadores en Psicología del Mercosur. 2006; Buenos Aires: Universidad de Buenos Aires. p. 0-2.
- Pureza JR, Gonçalves HA, Branco L, Grassi-Oliveira R, Fonseca RP. Executive functions in late childhood: age differences among groups. *Psychol Neurosci.* 2013;6:79-88.
- Robinson S, Goddard L, Dritschel B, Wisley M, Howlin P. Executive functions in children with autism spectrum disorders. *Brain Cogn.* 2009;71:362-8.
- Abusamra V, Raiter A, Ferrere A. Una perspectiva cognitiva en el estudio de la comprensión de textos. *Psico.* 2008;39:352-61.
- Borella E, Carretti B, Pelegrina S. The specific role of inhibition in reading comprehension in good and poor comprehenders. *J Learn Disabil.* 2010;43:541-52.
- White SJ, Burgess PW, Hill EL. Impairments on "open-ended" executive function tests in Autism. *Autism Res.* 2009;2:138-47.
- Ames CS, White SJ. Brief report: Are ADHD traits dissociable from the autistic profile? Links between cognition and behaviour. *J Autism Dev Disord.* 2011;41:357-63.
- Bayliss DM, Roodenrys S. Executive processing and attention deficit hyperactivity disorder: an application of the supervisory attentional system. *Dev Neuropsychol.* 2000;17:161-80.
- Clark C, Prior M, Kinsella GJ. Do executive function deficits differentiate between adolescents with ADHD and oppositional defiant/conduct disorder? A neuropsychological study using the Six Elements Test and Hayling Sentence Completion Test. *J Abnorm Child Psychol.* 2000;28:403-14.
- Campbell C, Barrett S, Shannon C, Hoy K, Rushe T, Cooper S, et al. The relationship between childhood trauma and neuropsychological functioning in first episode psychosis. *Psychosis.* 2013;5:48-59.
- Channon S, Pratt P, Robertson MM. Executive function, memory, and learning in Tourette's syndrome. *Neuropsychology.* 2003;17:247-54.
- Stedal K, Rose M, Frampton I, Landroø NI, Lask B. The neuropsychological profile of children, adolescents, and young adults with anorexia nervosa. *Arch Clin Neuropsychol.* 2012;27:329-37.
- Pureza JR, Jacobsen GM, Oliveira RG, Fonseca RP. Relationships between executive functions tasks in late childhood. *Psychol Neurosci.* 2011;4:369-76.
- Siqueira L de S, Scherer LC, Reppold CT, Fonseca RP. Hayling test - adult version: applicability in the assessment of executive functions in children. *Psychol Neurosci.* 2010;3:189-94.
- Forns J, Esnaola M, López-Vicente M, Suades-González E, Alvarez-Pedrerol M, Julvez J, et al. The n-back test and the attentional network task as measures of child neuropsychological development in epidemiological studies. *Neuropsychology.* 2014;28:519-29.
- Packwood S, Hodgetts HM, Tremblay S. A multiperspective approach to the conceptualization of executive functions. *J Clin Exp Neuropsychol.* 2011;33:456-70.
- Rabelo ISA, Pacanaro SV, Rossetti M de O, de Sá Leme IFA, de Castro NR, Güntert CM, et al. Color trails test: a Brazilian normative sample. *Psychol Neurosci.* 2010;3:93-9.
- Burgess PW, Shallice T. Response suppression, initiation, and strategy use following frontal lobe lesions. *Neuropsychologia.* 1996;34:263-73.
- Fonseca RP, Jacobsen GM, Pureza JR. O que um bom teste neuropsicológico deve ter? In: de Salles JF, Haase VG, Malloy-

- Diniz LF, editores. Neuropsicologia do desenvolvimento: infância e adolescência. Porto Alegre: Artmed; 2015.
37. Barbosa GA, Gouveia V V. O fator hiperatividade do Questionário de Conners: validação conceptual e normas diagnósticas. *Temas Teor Prat Psiquiatr.* 1993;23:188-202.
  38. Angelini AL, Alves ICB, Custódio EM, Duarte WF, Duarte JLM. Manual matrizes progressivas coloridas de Raven: escala especial. São Paulo: Centro Editor de Testes e Pesquisas em Psicologia; 1999.
  39. Fonseca RP, Casarin FS, Oliveira Crde, Gindri G, Ishigaki ECS, Ortiz KZ, et al. Adaptação de instrumentos neuropsicológicos verbais: um fluxograma de procedimentos para além da tradução. *Interação Psicol.* 2011;15:59-69.
  40. Fagundes AJ. Descrição, definição e registro de comportamento. São Paulo: EDICON; 1985.
  41. Andrés AM, Marzo PF. Delta: a new measure of agreement between two raters. *Br J Math Stat Psychol.* 2004;57:1-19.
  42. Heilmann J, Miller JF, Nockerts A, Dunaway C. Properties of the narrative scoring scheme using narrative retells in young school-age children. *Am J Speech Lang Pathol.* 2010;19:154-66.
  43. Nampijja M, Apule B, Lule S, Akurut H, Muhangi L, Elliott AM, et al. Adaptation of Western measures of cognition for assessing 5-year-old semi-urban Ugandan children. *Br J Educ Psychol.* 2010;80:15-30.
  44. Argollo N, Bueno OFA, Shayer B, Godinho K, Abreu K, Durán P, et al. Adaptação transcultural da Bateria NEPSY - avaliação neuropsicológica do desenvolvimento: estudo-piloto. *Aval Psicol.* 2009;8:59-75.
  45. Gutierrez-Clellen VF, Quinn R. Assessing narratives of children from diverse cultural/linguistic groups. *Lang Speech Hear Serv Sch.* 1993;24:2-9.
  46. Radanovic M, Mansur LL. Performance of a Brazilian population sample in the Boston Diagnostic Aphasia Examination. a pilot study. *Braz J Med Biol Res.* 2002;35:305-17.
  47. Altmann GTM. The language machine: psycholinguistics in review. *Br J Psychol.* 2001;92:129-70.
  48. Berman RA. The psycholinguistics of developing text construction. *J Child Lang.* 2008;35:735-71.
  49. Biaggio AMB, Barreto MSL. Adaptação brasileira de uma medida objetiva de julgamento moral. *Arq Bras Psicol.* 1991;43:107-19.
  50. Acevedo A, Krueger KR, Navarro E, Ortiz F, Manly JJ, Padilla-Vélez MM, et al. The Spanish translation and adaptation of the Uniform Data Set of the National Institute on Aging Alzheimer's Disease Centers. *Alzheimer Dis Assoc Disord.* 2009;23:102-9.
  51. Casarin FS, Scherer LC, Ferré P, Ska B, Parente MAP de M, Joannette Y, et al. Adaptação do protocolo MEC de Poche e da bateria MAC expandida: bateria MAC Breve. *Psico.* 2013;44:288-99.
  52. Wilson BA. Avanços recentes em reabilitação neuropsicológica. Neuropsicologia e as interfaces com as neurociências. São Paulo: Casa do Psicólogo; 2007.
  53. Janczura GA, Castilho GM, Rocha NO, van Erven TJC, Huang TP. Normas de concreude para 909 palavras da língua portuguesa. *Psic Teor e Pesq.* 2007;23:195-203.
  54. Stein LM, Gomes CFA. Normas brasileiras para listas de palavras associadas: associação semântica, concreude, frequência e emocionalidade. *Psic Teor e Pesq.* 2009;25:537-46.
  55. Nascimento E, Figueiredo VLM. WISC-III e WAIS-III: alterações nas versões originais americanas decorrentes das adaptações para uso no Brasil. *Psicol Reflex Crit.* 2002;15:603-12.
  56. Pasquali L. Testes referentes a construto: teoria e modelo de construção. In: Pasquali L. Instrumentação psicológica fundamentos e práticas. Porto Alegre: Artmed; 2010. p. 165-98.
  57. Bandeira MB, Felício CM, Cesari L. Percepção de mudanças pelos familiares como resultado do tratamento em serviços de saúde mental: validação da escala EMP-familiares. *Rev Bras Psiquiatr.* 2010;32:283-7.
  58. Nogueira GJ, Castro A, Naveira L, Nogueira-Antuñano F, Natinzon A, Gigli SL, et al. [Evaluation of the higher brain functions in 1st and 7th grade schoolchildren belonging to two different socioeconomic groups]. *Rev Neurol.* 2005;40:397-406.
  59. Rosselli-Cock M, Matute-Villaseñor E, Ardila-Ardila A, Botero-Gómez VE, Tangarife-Salazar GA, Echevarría-Pulido SE, et al. [Neuropsychological Assessment of Children: a test battery for children between 5 and 16 years of age. A Colombian normative study]. *Rev Neurol.* 2004;38:720-31.
  60. Casarin FS, Wong CEI, Parente MAMP, Salles JF, Fonseca R. Comparação de desempenho neuropsicológico entre estudantes de escolas públicas e privadas brasileiras. *Span J Psychol.* 2012;15:942-51.
  61. Dufouil C, Alperovitch A, Tzourio C. Influence of education on the relationship between white matter lesions and cognition. *Neurology.* 2003;60:831-6.

**Correspondence:**

Larissa de Souza Siqueira  
Av. Ipiranga, 6681, prédio 11, 9º andar, sala 932  
90619-900 - Porto Alegre, RS - Brazil  
Tel./Fax: +55 (51) 3320.3500 - Ext. 7742  
E-mail: larissassiqueira@gmail.com