



# Cultural adaptation and reliability assessment of the Hammersmith neonatal neurological examination for Brazilian newborns at risk of cerebral palsy

## *Adaptação cultural e avaliação da confiabilidade do exame neurológico neonatal de Hammersmith para recém-nascidos brasileiros com risco de paralisia cerebral*

Mayara Thais Correr<sup>1</sup> Luzia Lara Pfeifer<sup>1</sup>

<sup>1</sup> Universidade de São Paulo, Faculdade de Medicina de Ribeirão Preto, Departamento de Neurociências e Ciências do Comportamento, Programa de Pós-graduação em Neurologia, Ribeirão Preto SP, Brazil.

Address for correspondence Luzia Lara Pfeifer (email: luziara@fmrp.usp.br).

Arq. Neuropsiquiatr. 2023;81(1):47–54.

### Abstract

**Background** Reliable instruments that lead to early diagnosis for CP are extremely important so that these children are referred for early stimulation, benefiting their development.

**Objective** To perform a cross-cultural adaptation and reliability assessment of a Brazilian version of the Hammersmith Neonatal Neurological Examination (HNNE), expanded and summarized.

**Methods** A methodological, cross-sectional, nonexperimental quantitative analysis was conducted in two phases as follows: cultural adaptation of the HNNE, expanded and summarized, and reliability assessment of the Brazilian version of the HNNE. Phase one was developed in five stages (initial translation, synthesis of the translation, a committee of experts, backtranslation, and submission to the author), with the semantic questions, content, and face validity being evaluated. Phase two included 143 newborns and we analyzed the internal consistency, stability, and equivalence (intra- and interexaminer) of the instrument. Internal consistency was calculated using Cronbach's alpha, and intra- and interexaminer reliability and reproducibility assessed through test-retest were calculated using the intraclass correlation coefficient.

**Results** Although internal consistency, assessed using Cronbach's alpha, showed unsatisfactory results, the results of inter- and intraexaminer equivalence showed a high agreement between the evaluations in all domains. The test-retest also showed excellent agreement between the domains.

**Conclusions** The Brazilian HNNE expanded and summarized versions can be considered to be adapted and reliable for the neurological assessment of Brazilian newborns.

### Keywords

- ▶ Cerebral Palsy
- ▶ Early Diagnosis
- ▶ Neurology

received  
November 18, 2021  
received in its final form  
February 21, 2022  
accepted  
May 8, 2022

DOI <https://doi.org/10.1055/s-0042-1758863>.  
ISSN 0004-282X.

© 2023. Academia Brasileira de Neurologia. All rights reserved. This is an open access article published by Thieme under the terms of the Creative Commons Attribution 4.0 International License, permitting copying and reproduction so long as the original work is given appropriate credit (<https://creativecommons.org/licenses/by/4.0/>).  
Thieme Revinter Publicações Ltda., Rua do Matoso 170, Rio de Janeiro, RJ, CEP 20270-135, Brazil

to identify changes in neurological development and early referral to the stimulation or early rehabilitation units and as a promising option to be used in the context of primary care in Brazil.

## Resumo

**Antecedentes** As avaliações neurológicas que levam ao diagnóstico precoce permitem o acesso oportuno à intervenção em um período em que os maiores ganhos são possíveis devido à neuroplasticidade.

**Objetivos** Realizar a adaptação transcultural e avaliação da confiabilidade da versão brasileira do Hammersmith Neonatal Neurological Examination (HNNE), ampliada e resumida.

**Métodos** Foi realizada análise quantitativa metodológica, transversal e não experimental em duas fases: adaptação cultural do HNNE, ampliada e resumida, e avaliação da confiabilidade da versão brasileira do HNNE. A primeira fase foi desenvolvida em cinco etapas (tradução inicial, síntese da tradução, comitê de especialistas, retrotradução e submissão ao autor), sendo avaliadas as questões semânticas, conteúdo e validade de face. A fase dois incluiu 143 recém-nascidos e foram analisadas a consistência interna, estabilidade e equivalência (intra e interexaminador) do instrumento. A consistência interna foi calculada pelo alfa de Cronbach, e a confiabilidade e reprodutibilidade intra e interexaminadores avaliadas por meio do teste-reteste foram calculadas pelo coeficiente de correlação intraclassa.

**Resultados** Embora a consistência interna, avaliada pelo alfa de Cronbach, tenha apresentado resultados insatisfatórios, os resultados da equivalência inter e intraexaminadores mostraram alta concordância entre as avaliações em todos os domínios. O teste-reteste também apresentou excelente concordância entre os domínios.

**Conclusões** As versões brasileiras ampliadas e resumidas do HNNE podem ser consideradas adaptadas e confiáveis para avaliação neurológica de recém-nascidos brasileiros por identificar alterações no desenvolvimento neurológico e encaminhamento precoce para unidades de estimulação ou reabilitação precoce e como uma opção promissora para uso no contexto da atenção básica no Brasil.

## Palavras-chave

- ▶ Paralisia Cerebral
- ▶ Diagnóstico Precoce
- ▶ Neurologia

## INTRODUCTION

Neurological assessment is one of the most widely used clinical tools to monitor the development of babies at risk of neurological disabilities.<sup>1,2</sup> Currently, most scientific studies are linked to the assessment of global development and the acquisition of motor, cognitive, language, and socioemotional skills.<sup>3</sup> However, studies on neurological assessments that lead to early diagnosis are essential, allow timely access to intervention in a period where the greatest gains are possible due to neuroplasticity. Late diagnosis can be detrimental to the development of a child and can deprive them of early intervention for months or even years.<sup>4</sup>

The tools with the best predictive validity to detect cerebral palsy (CP) are neonatal magnetic resonance imaging, Prechtl Qualitative Assessment of General Movements, and the Hammersmith Neurological Examination.<sup>1</sup> The Hammersmith Neurological Examination has two versions: the expanded and summarized versions of the Hammersmith Neonatal Neurological Examination (HNNE), which evaluates newborns (NBs) up to 3 months old, and the

Hammersmith Infant Neurological Examination (HINE), which evaluates infants between 30 days and 24 months old.<sup>5</sup>

The expanded version of the HNNE consists of 34 items subdivided into 6 categories as follows: 1) posture and tone, 2) tone patterns, 3) reflexes, 4) movements, 5) abnormal signs, and 6) orientation and behavior. The examination takes ~ 15 minutes and can be used to assess NBs with unstable conditions, without the need to follow the sequence proposed in the evaluation form. One can also choose the most appropriate sequence in relation to the positioning of the baby or their alertness.<sup>6</sup> Each item has figures and descriptions and may be scored as 0.0 (abnormal), 0.5 (intermediate), or 1.0 (normal). The total score is calculated as the sum of the individual item scores, with the normal range being 30.5 to 34. If the global score is in the borderline zone, it does not necessarily mean that the assessed NB presents neurological abnormalities, but it identifies that regular neurological follow-up must be maintained.<sup>6</sup>

In 2005, a short and simplified version of the HNNE was prepared for screening, which consisted of 25 items under





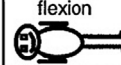




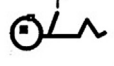
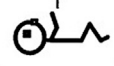


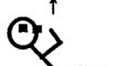
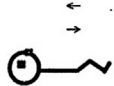
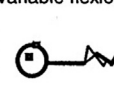
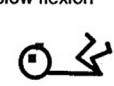

Hammersmith Neonatal Neurological Examination						CODE _____	D.O.E. _____	S T A T E	A S Y M M	
NAME _____		SEX _____	RACE _____	D.O.B. _____	AGE _____	G.A. _____	BW _____			
<b>Posture and tone</b>										
<b>POSTURE</b> Infant supine. Look mainly at position of legs but also note arms. <i>Score predominant posture.</i>	arms & legs extended or very slightly flexed	Legs slightly flexed	legs well flexed but not adducted	legs well flexed & adducted near abdomen	abnormal posture: a) opisthotonus b) marked leg extension, strong arm flexion					
<b>ARM RECOIL</b> Take both hands, quickly extend arms parallel to the body, Count to three. Release. Repeat 3 times.	arms do not flex	arms flex slowly, not always; not completely	arms flex slowly; more completely	arms flex quickly and completely	arms difficult to extend; snap back forcefully					
<b>ARM TRACTION</b> Hold wrist and pull arm upwards. Note flexion at elbow and resistance while shoulder lifts off table. <i>Test each side separately.</i>	arms remain straight; no resistance felt	arms flex slightly or some resistance felt	arms flex well till shoulder lifts, then straighten	arms flex at approx 100° & maintained as shoulder lifts	flexion of arms <100°; maintained when body lifts up	R  L	R  L	R  L	R  L	R  L
<b>LEG RECOIL</b> Take both ankles in one hand, flex hips + knees. Quickly extend. Release. Repeat 3 times.	No flexion	incomplete or variable flexion	complete but slow flexion	complete fast flexion	legs difficult to extend; snap back forcefully					

Figure 1 A short and simplified version of the HNNE, the behaviors listed in the first and last columns are abnormal.

the following categories: posture and tone, movements, reflexes, guidance, and behavior.<sup>5</sup> The behaviors listed in the first and last columns are abnormal for term infants (► Figure 1); therefore, if two or more items in these columns are scored or one or more of the abnormal signs that are listed at the end of the instrument are noted, then the infant must be evaluated by the full version.<sup>7</sup> Both versions can be applied to premature infants.<sup>6</sup>

In Brazil, validated assessment tools for predicting neurological disorders are rare. The HNNE has been identified as one of the best and simplest neurological examinations for the early diagnosis of neurological impairment in low- and high-risk neonates, and it is an easily applied tool, even by inexperienced professionals.<sup>2,8</sup> Although there are already translations in Portuguese from Brazil and Portugal, elaborated by Tathiana Ghisi de Souza e Moyra Aloia Romero (linked in hammersmith-neuro-exam.com), this translation was done freely, without the cross-adaptation process. The cultural adaptation process includes a translation by two independent translators, the committee of experts, synthesis of translations, and backtranslation.<sup>9,10</sup> This way, we decided to perform a cross-cultural adaptation and assess the reliability of a Brazilian version of the HNNE (expanded and summarized).

**METHODS**

The present methodological, cross-sectional, nonexperimental study, approved by the ethics committee under opinion No. 1,809,858, was conducted in 2 phases as follows:

cultural adaptation of the HNNE (expanded and summarized) and reliability assessment of the Brazilian version of the HNNE (expanded).

**Phase 1**

Cultural adaptation of the HNNE - expanded and summarized<sup>6</sup> to Brazilian neonates.

We obtained authorization from the authors of the instrument via email to carry out the HNNE translation, cultural adaptation, and psychometric validation of the expanded and summarized versions. The translation and cultural adaptation procedures followed the guidelines proposed by Wild et al.,<sup>9</sup> Beaton et al.,<sup>10</sup> Pasquali,<sup>11</sup> and Ferrer et al.,<sup>12</sup> in five stages as follows:

**Stage 1 - initial translation**

Two Brazilian translators (T1 and T2) independently translated the original version of the HNNE (expanded and summarized). T1 had knowledge in the field of neuropediatrics and knew the concepts examined by the instrument, while T2 had none of this knowledge. However, both translators had proficiency in both languages (English and Portuguese) and prepared translation versions 1 and 2 (VT1 and VT2), respectively.

**Stage 2 - synthesis of the translations**

A technical committee comprising an experienced researcher in the field of neuropediatrics, a neurology postgraduate, and a doctor-researcher in the field of neuropediatrics with >25 years of experience, all with proficiency in both

languages, was formed to compare VT1 and VT2 and to elaborate the Portuguese synthesized version (VSP). This procedure followed the recommendations made by Koller et al.<sup>13</sup> when we merged VT1 and VT2 with modifications/additions or used VT2 adapted to VT1 and vice-versa.

### **Stage 3 - committee of experts**

The original and VSP versions were analyzed for their semantic and content by a committee of experts comprising 10 professionals with experience in neuropediatrics. Face validity was also performed at this stage, since the committee was formed by professionals in the area, to point out clarity in the items that made up the instrument. The technical committee analyzed all suggestions of experts and re-evaluated and restructured all items with < 80% agreement and sent them to experts again until an acceptable deal was reached, as proposed by the literature.<sup>10,11</sup> Thus, the technical committee elaborated on the Portuguese consensual version (VCP), which was then sent for backtranslation.

### **Stage 4 - back-translation**

The VCP was backtranslated into English by two independent backtranslators who were native English speakers and had proficiency in Portuguese, both without knowledge of the original version of the instrument and without experience in neuropediatrics. The same technical committee of stage 2 analyzed the semantic, idiomatic, cultural, and conceptual equivalences and produced the synthesis of backtranslations 1 and 2 by elaborating the consensual back-translation version (VCRT). This step was important to validate whether the translation reflected the same content as the items in the original instrument.

### **Stage 5 - submission to the author**

The VCP and VCRT versions were sent to the author of the instrument and the team who approved the versions consolidating the final version (HNNE-Br) of the HNNE instrument in all their versions (HNNE expanded and summarized + HINE) for the Brazilian population. It is important to note that the authors did not make suggestions or indicate changes to be made in the documents sent to them.

The expanded HNNE had 34 items, each with 5 possible answers; thus, in the stages of cultural adaptation, it was broken down into 170 topics of analysis, in addition to items that included comments, observations, characterization of the evaluated items, and a summary of the instrument, totaling 214 topics translated and adapted to the Brazilian culture.

The summarized HNNE had 25 items, each with 5 possible answers, and it was broken down, totaling 157 topics translated and adapted to the Brazilian culture.

## **Phase 2**

Reliability of the Brazilian version of the HNNE (HNNE-Br) – expanded and summarized.

A convenience sample of 143 premature and full-term neonates who presented some risk for CP were recruited

from obstetric outpatient, neonatal maternity, and intermediate care units (UCIN), of two hospitals of a city located in the interior of the state of São Paulo, Brazil.

We analyzed the internal consistency, stability, and equivalence to verify the reliability of the HNNE-Br (expanded). The 25 items of the HNNE-Br summarized are present in the expanded version, so the results are the same for both versions:

- The Cronbach alpha test analyzed the internal consistency<sup>14</sup> of the HNNE-Br (expanded) and was performed with a total sample of 143 NBs. The evaluations took place as near as possible after birth, and some NBs were evaluated a few hours after birth, while others a few days later.
- Intra- and interexaminer reliability was analyzed to assess equivalence. The evaluations were filmed in the UCIN and maternity ward for scoring later. The footage was captured with the help of a collaborator who positioned the camera appropriately to record the assessment of each item of the evaluation. Participants could either be in an incubator (under intermediate or intensive care) or in the nursery. The first evaluation took place in loco and the second evaluation took place 14 days after the first, through the images assisted by evaluator 1 to prevent his memory from influencing the results. To achieve inter-examiner reliability, another health professional with neonatal experience and training in occupational therapy was invited to participate in the research (evaluator 2).
- The stability of the instrument was assessed by test-retest reliability with 30 babies as a sample. The test-retest was performed by evaluator 1; however, between these evaluations, the 14-day interval was not respected because the neonates of the units (maternity and UCIN) could be discharged at any time, making it difficult for them to return to the data collection units. In addition, there was a great variability in the behavior of the newborn, exponentially changing the results of the evaluation. Thus, the second evaluation took place on the same day as the first with a sleep interval, breastfeeding, or a routine examination at the unit. The criterion used between the evaluations was to ensure that the behavioral state of the NB in the second evaluation was similar to the behavioral state as observed in the first evaluation.

## **Data analysis**

Data were analyzed using IBM SPSS Statistics for Windows version 22.0 (IBM Corp., Armonk, NY, USA). The participants received identification numbers to maintain anonymity and ensure blinding of the researcher responsible for the data analysis. Internal consistency was calculated using Cronbach alpha, which has recognized limits between 0.70 and 0.90.<sup>15</sup> The intra- and interexaminer reliability and reproducibility assessed through test-retest were calculated using the intra-class correlation coefficient (ICC 3.1). Each rater evaluated each subject, and the reliability was calculated from a single measurement. The ICC values interpretation considered correlations < 0.41 - weak, between 0.41 and 0.60 - moderate; between 0.61 and 0.80 - strong or substantial, and between 0.81 and 1.00 - almost perfect.<sup>16</sup>

## RESULTS

After analyzing the translations from English to Portuguese (PV1 and PV2) of the 214 topics that structured the expanded instrument, > 80% agreement was found between the original and synthesized versions for 201 items. Eight items obtained 70% agreement, 4 obtained 60%, and only 1 item, which was the instrument's title, obtained agreement below 50%. Only 1 person agreed with the title (equivalent to 10% agreement); therefore, the title was changed.

In the equivalence assessment of the summarized HNNE, 25 items were evaluated, and 157 topics were translated and adapted to Brazilian culture. Nine specialists participated in this process, and there was an agreement of > 80% between the original and synthesized versions in 141 items. Thirteen items obtained 77.7% agreement, 1 obtained 66.6%, and 2 < 50%. Regarding the semantic validity, idiomatic validity, and conceptual validity, 3, 8, and 28 items, respectively, were changed.

The technical committee restructured the items with the lower-than-expected agreement, created the PCV, and subsequently sent it for backtranslation.

In stage 5, we sent the PCV, the RTCV, and the original version to the authors so that they could assess the equivalence of the versions and, if necessary, suggest some modifications. However, no changes were requested, or suggestions were added; thus, we retained the versions as presented. In this way, we consolidated each the final versions of each instrument (HNNE - Brazilian expanded and summarized versions [HNNE-Br]).

In phase 2 of reliability, 143 neonates participated in the study, with 98 (68.5%) full-term and 45 (31.5%) preterm babies. There were no extreme preterm births in this sample (gestational age [GA] < 28 weeks); however, 21 intermediate premature infants (GA between 28 and 34 weeks) and 24 late preterm infants (GA between 34 and 36 weeks) participated in the present study.

Unsatisfactory results were found regarding the internal consistency of each item, which was assessed using Cronbach alpha. These can be seen in ► **Table 1**.

The interexaminer equivalence was performed by two different evaluators, with diverse academic backgrounds (physiotherapy and occupational therapy) in a subgroup formed by 30 newborns. The ICCs revealed high agreement between the evaluations in all domains (► **Table 2**).

Intraexaminer reliability was performed in a subgroup of 30 newborns with a 14-day interval between onsite assessments and filming assessments, according to the interval proposed by Terwee et al.<sup>17</sup> It was found that there was a high level of agreement by domain between the evaluations, which can be seen in ► **Table 2**.

The test-retest was applied to assess the measurement stability and calculate the ICC, which showed an excellent agreement between the domains, as shown in ► **Table 3**.

## DISCUSSION

The cultural adaptation of the HNNE-Br instruments (expanded and summarized versions) to Brazilian Portuguese

**Table 1** Internal consistency by Hammersmith Neonatal Neurological Examination items

Item	Cronbach alpha $\alpha$
Posture	0.555
Arm recoil	0.543
Arm traction	0.564
Leg recoil	0.537
Leg traction	0.554
Popliteal angle	0.527
Head control (1)	0.538
Head control (2)	0.525
Head lag	0.537
Ventral suspension	0.549
Flexor tone (1)	0.564
Flexor tone (2)	0.542
Leg extensor tone	0.588
Neck extensor tone	0.571
Increased extensor tone	0.557
Tendon reflex	0.542
Suck/gag	0.569
Palmar grasp	0.551
Plantar grasp	0.548
Placing	0.539
Moro reflex	0.546
Spontaneous movement (quantity)	0.515
Spontaneous movement (quality)	0.516
Head raising prone	0.553
Abnormal hand or toes postures	0.545
Tremor	0.549
Startle	0.539
Eye appearances	0.543
Auditory orientation	0.536
Visual orientation	0.541
Alertness	0.537
Irritability	0.530
Consolability	0.529
Cry	0.527

was carried out according to internationally recognized procedures, following the guidelines proposed by Ferrer et al.,<sup>12</sup> Beaton et al.<sup>10</sup> and Guillemin et al.<sup>18</sup>

The cultural adaptation process was initiated by translating the instruments into the Brazilian Portuguese language through two independent and bilingual translators,<sup>10</sup> with different academic backgrounds to increase the probability of finding more relevant terms.<sup>19</sup> According to Tanzer,<sup>19</sup> an adequate translation must promote linguistic, cultural, and contextual understanding so that the evaluated construct

**Table 2** Interexaminer and intraexaminer reliability by intraclass correlation tests

Domain	Interexaminer		Intraexaminer			
	ICC	95%CI	p-value	ICC	95%CI	-p-value
Posture and tone	0.964	0.924–0.983	< 0.001	0.944	0.887–0.973	< 0.001
Tone patterns	0.969	0.934–0.985	< 0.001	1.000	–	< 0.001
Reflexes	0.951	0.897–0.977	< 0.001	0.962	0.923–0.982	< 0.001
Movements	0.972	0.942–0.987	< 0.001	0.993	0.986–0.997	< 0.001
Abnormal signs/patterns	0.912	0.816–0.958	< 0.001	0.990	0.979–0.995	< 0.001
Orientation and behavior	0.841	0.667–0.924	< 0.001	0.964	0.925–0.983	< 0.001
Total	0.964	0.925–0.983	< 0.001	0.955	0.908–0.978	< 0.001

Abbreviations: CI, confidence interval; ICC, intraclass correlation coefficient.

**Table 3** Test-retest reliability by intraclass correlation

Domain	ICC	95%CI	p-value
Posture and tone	0.996	0.991–0.998	< 0.001
Tone patterns	0.981	0.960–0.991	< 0.001
Reflexes	0.980	0.957–0.990	< 0.001
Movements	0.961	0.918–0.981	< 0.001
Abnormal signs/patterns	0.995	0.990–0.998	< 0.001
Orientation and behavior	0.854	0.693–0.930	< 0.001
Total	0.983	0.964–0.992	< 0.001

Abbreviations: CI, confidence interval; ICC, intraclass correlation coefficient.

can maintain its integrality in both cultures. On the other hand, poorly translated instruments interfere negatively with data validity.<sup>9</sup>

Subsequently, the technical committee compared the translated versions to develop a single and synthesized version,<sup>20</sup> which was sent to the expert committee, together with the original version to be evaluated regarding semantic, idiomatic, conceptual, cultural, and experiential discrepancies. A multidisciplinary committee of experts was essential to ensure the accuracy of the content.<sup>21</sup>

Specialists must judge the relevance and range of the items of each instrument to assess the pertinence and scope, including face validity.<sup>22</sup> The experience of the experts is crucial to preserve the characteristics of the instrument that make it appropriate for the target population.<sup>19</sup> In this case, all suggestions made by the expert committee were rigorously examined and considered in the consensual versions of the instruments by obtaining satisfactory levels of agreement.

Ferrer et al.<sup>12</sup> point out that after obtaining the translations, they should be evaluated by a committee of experts and only then one can proceed with the backtranslation. Backtranslation is crucial because it allows authors to assess whether they maintain the same conceptual idea as the original instrument, thus preserving its construct.<sup>20</sup> The authors of the instrument agreed with the HNNE-Br versions presented, without any objection, which consolidated the final version without further changes.

The preliminary validation of the HNNE-Br, culturally adapted to Brazilian Portuguese, was carried out according to psychometric procedures recognized in the scientific literature by analyzing internal consistency and reproducibility.<sup>20,23</sup>

In traditional psychometry, 10 respondents are recommended for each questionnaire item. In the expanded version of the HNNE, there are 34 items, while in the shorter one, there are 25 items. Following this concept, we would need to recruit between 250 and 340 participants. However, Sapnas et al.<sup>24</sup> demonstrated that subsamples with between 50 and 100 subjects are sufficient to analyze the psychometric properties, especially social constructs. They believed that 10 respondents per item are more than necessary and guaranteed that 100 subjects are sufficient to verify the initial psychometric properties of an instrument being tested in another population, allowing a desirable conclusion to be reached.<sup>24</sup> Therefore, the sample size of the present study was adequate.

Cronbach alpha has recognized limits between 0.70 and 0.90, with values < 0.70 indicating nonconsistency.<sup>25</sup> Although the HNNE-Br Cronbach's alpha coefficient values are below what are considered acceptable, studies on the psychometric properties of the HNNE in its original version or other versions have not been found, making a comparison impossible. Keszei et al.<sup>26</sup> reported that low internal consistency values can mean that the items measure different attributes. It happens on the HNNE because this instrument

evaluates the motor, visual, and auditory systems, and a baby with CP may have an impaired motor system while the other systems may remain intact, corroborating the lack of consistency between the items.

The inter-rater reliability, performed in a subgroup formed by 30 NBs through the initial filmed evaluations, showed high agreement, agreeing with the results provided by the authors when developing the instrument<sup>27</sup> and by Eeles et al.,<sup>28</sup> who found excellent reliability ( $ICC > 0.74$ ) for the total score when evaluating premature NBs.

Evaluating through filming is widely used in the literature for intrarater reliability, being considered by many to be a satisfactory method.<sup>29,30</sup> However, some researchers highlight difficulties in evaluating individuals from videos, as filming limits viewing from different angles.<sup>31,32</sup> Thus, we invited one occupational therapist, who worked in the intermediate care unit, to carry out the filming, to capture the details of the evaluation, and highlight the response of the NB after a stimulus, such as the reflex. This form of assessment allows no changes between the behavior or response of the individual, which could be different if the reassessment occurred after 14 days, thus, compromising data reliability.

As with previous measurements, the period between the test and retest is a factor that must be considered. If carried out in too short a period, the results may be contaminated by the memory effect of the first application. On the other hand, if carried out after a long period, results are susceptible to changes and the acquisition of new skills and may compromise the interpretation of the obtained reliability coefficient.<sup>26</sup>

The test-retest evaluations took place on the same day because the acquisitions and behaviors of the neonates were quite variable. The results showed almost perfect agreement in nearly all items, with the lowest values attributed to the behavior domain (which may vary if the newborn is hungry, sleepy, colic, or even due to the high degree of handling in the sector for collections of evaluations). The highest values were for posture and tone, which, despite being variable, did not change much on the same day.

In conclusion, the analyzed properties suggest that the HNNE is adapted and reliable for the Brazilian population and can be a useful instrument for the neurological assessment of Brazilian newborns, to identify changes in neurological development, and to refer them early to the stimulation or early rehabilitation units. The HNNE is an easy and fast application tool, has open access, and, therefore, is a promising option to be used in the context of primary care in Brazil.

The availability of an instrument with these characteristics will favor the inclusion of child development surveillance in the daily lives of health professionals, enabling intervention at an appropriate time in cases suspected of alterations. In addition, using a standardized and reliable instrument for Brazil will also allow the generation of epidemiological information to promote public policies and to compare with studies carried out in other countries.

#### Authors' Contributions

MTC: main researcher; LIP: researcher and advisor.

#### Conflict of Interest

The authors have no conflict of interests to declare.

#### References

- Novak I, Hines M, Goldsmith S, Barclay R. Clinical prognostic messages from a systematic review on cerebral palsy. *Pediatrics* 2012;130(05):e1285–e1312. Doi: 10.1542/peds.2012-0924
- Romeo DM, Ricci D, Brogna C, Mercuri E. Use of the Hammersmith Infant Neurological Examination in infants with cerebral palsy: a critical review of the literature. *Dev Med Child Neurol* 2016;58(03):240–245. Doi: 10.1111/dmcn.12876
- Bodeau-Livinec F, Zeitlin J, Blondel B, et al; Etude Epidemiologique sur les Petits Ages Gestationnels (EPIPAGE) group. Do very preterm twins and singletons differ in their neurodevelopment at 5 years of age? *Arch Dis Child Fetal Neonatal Ed* 2013;98(06):F480–F487. Doi: 10.1136/archdischild-2013-303737
- Bosanquet M, Copeland L, Ware R, Boyd R. A systematic review of tests to predict cerebral palsy in young children. *Dev Med Child Neurol* 2013;55(05):418–426. Doi: 10.1111/dmcn.12140
- Mercuri E, Ricci D, Pane M, Baranello G. The neurological examination of the newborn baby. *Early Hum Dev* 2005;81(12):947–956. Doi: 10.1016/j.earlhumdev.2005.10.007
- Dubowitz L, Dubowitz V, Mercuri E. *The neurological assessment of the preterm and full term infant*. Clinics in Developmental Medicine. Vol. 148. London: McKeith Press; 1999
- Dubowitz L, Ricci D, Mercuri E. The Dubowitz neurological examination of the full-term newborn. *Ment Retard Dev Disabil Res Rev* 2005;11(01):52–60. Doi: 10.1002/mrdd.20048
- Novak I, Morgan C, Adde L, et al. Early, Accurate Diagnosis and Early Intervention in Cerebral Palsy: Advances in Diagnosis and Treatment. *JAMA Pediatr* 2017;171(09):897–907. Doi: 10.1001/jamapediatrics.2017.1689
- Wild D, Grove A, Martin M, et al; ISPOR Task Force for Translation and Cultural Adaptation. Principles of Good Practice for the Translation and Cultural Adaptation Process for Patient-Reported Outcomes (PRO) Measures: report of the ISPOR Task Force for Translation and Cultural Adaptation. *Value Health* 2005;8(02):94–104. Doi: 10.1111/j.1524
- Beaton DE, Bombardier C, Guillemin F, Ferraz MB. Guidelines for the process of cross-cultural adaptation of self-report measures. *Spine* 2000;25(24):3186–3191. Doi: 10.1097/00007632-200012150-00014
- Pasquali L. *Psicometria: teoria dos testes na psicologia e na educação*. Petrópolis: Editora Vozes; 2003
- Ferrer M, Alonso J, Prieto L, et al. Validity and reliability of the St George's Respiratory Questionnaire after adaptation to a different language and culture: the Spanish example. *Eur Respir J* 1996;9(06):1160–1166. Doi: 10.1183/09031936.96.09061160
- Koller M, Kantzer V, Mear I, et al; ISOQOL TCA-SIG. The process of reconciliation: evaluation of guidelines for translating quality-of-life questionnaires. *Expert Rev Pharmacoecon Outcomes Res* 2012;12(02):189–197. Doi: 10.1586/erp.11.102
- Cronbach LJ. Coefficient Alpha and the Internal Structure of Tests. *Psychometrika* 1951;16:297–334. Doi: 10.1007/BF02310555
- Nunnally JC. *Psychometric Theory*. 2nd ed. New York: McGraw Hill; 1978
- Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics* 1977;33(01):159–174
- Terwee CB, Bot SD, de Boer MR, et al. Quality criteria were proposed for measurement properties of health status questionnaires. *J Clin Epidemiol* 2007;60(01):34–42. Doi: 10.1016/j.jclinepi.2006.03.012

- 18 Guillemin F, Bombardier C, Beaton D. Cross-cultural adaptation of health-related quality of life measures: literature review and proposed guidelines. *J Clin Epidemiol* 1993;46(12):1417–1432. Doi: 10.1016/0895-4356(93)90142-n
- 19 Tanzer NK. Developing tests for use in multiple languages and cultures: A plea for simultaneous development. In: Hambleton RK, Merenda PF, Spielberger CD, eds. *Adapting educational and psychological tests for cross-cultural assessment*. Mahwah: Lawrence Erlbaum; 2005:235–64
- 20 Borsa JC, Damasio BF, Bandeira DR. Adaptação e validação de instrumentos psicológicos entre culturas: algumas considerações. *Paidéia* 2012;22(53):423–432. Doi: 10.1590/S0103-863x2012000300014
- 21 Epstein J, Santo RM, Guillemin F. A review of guidelines for cross-cultural adaptation of questionnaires could not bring out a consensus. *J Clin Epidemiol* 2015;68(04):435–441. Doi: 10.1016/j.jclinepi.2014.11.021
- 22 Mokkink LB, et al. *The COSMIN Manual*. Published 2012. Accessed October, 2015. Web site: <http://www.cosmin.nl/images/upload/files/COSMIN%20checklist%20manual%20v9.pdf>
- 23 Reichenheim ME, Moraes CL. Operacionalização de adaptação transcultural de instrumentos de aferição usados em epidemiologia. *Rev Saude Publica* 2007;41(04):665–673. Doi: 10.1590/S0034-89102006005000035
- 24 Sapnas KG, Zeller RA. Minimizing sample size when using exploratory factor analysis for measurement. *J Nurs Meas* 2002;10(02):135–154. Doi: 10.1891/jnum.10.2.135.52552
- 25 Nunnally JC. *Psychometric Theory*. 2nd ed. New York: McGraw Hill; 1978
- 26 Keszei AP, Novak M, Streiner DL. Introduction to health measurement scales. *J Psychosom Res* 2010;68(04):319–323. Doi: 10.1016/j.jpsychores.2010.01.006
- 27 Haataja L, Mercuri E, Regev R, et al. Optimality score for the neurologic examination of the infant at 12 and 18 months of age. *J Pediatr* 1999;135(2 Pt 1):153–161. Doi: 10.1016/s0022-3476(99)70016-8
- 28 Eeles AL, Olsen JE, Walsh JM, et al. Reliability of neurobehavioral assessments from birth to term equivalent age in preterm and term born infants. *Phys Occup Ther Pediatr* 2017;37(01):108–119. Doi: 10.3109/01942638.2015.1135845
- 29 Daum C, Gheorghita F, Spatola M, et al. Interobserver agreement and validity of bedside 'positive signs' for functional weakness, sensory and gait disorders in conversion disorder: a pilot study. *J Neurol Neurosurg Psychiatry* 2015;86(04):425–430. Doi: 10.1136/jnnp-2013-307381
- 30 Lowery JP, Hayes JR, Sis M, Griffith A, Taylor D. Pacific acuity test: testability, validity, and interobserver reliability. *Optom Vis Sci* 2014;91(01):76–85. Doi: 10.1097/OPX.0000000000000104
- 31 Wong CK. Interrater reliability of the Berg Balance Scale when used by clinicians of various experience levels to assess people with lower limb amputations. *Phys Ther* 2014;94(03):371–378. Doi: 10.2522/ptj.20130182
- 32 Rathke KM, Schäuble B, Fessler AJ, So EL. Reliability of seizure semiology in patients with 2 seizure foci. *Arch Neurol* 2011;68(06):775–778. Doi: 10.1001/archneurol.2011.97