Validity and reliability of the Persian ABILHAND-Kids in a sample of Iranian children with cerebral palsy

Elham Mohammadkhani-Pordanjani, Carlyne Arnould, Parvin Raji, Noureddin Nakhostin Ansari & Scott Hasson

To cite this article: Elham Mohammadkhani-Pordanjani, Carlyne Arnould, Parvin Raji, Noureddin Nakhostin Ansari & Scott Hasson (2019): Validity and reliability of the Persian ABILHAND-Kids in a sample of Iranian children with cerebral palsy, Disability and Rehabilitation, DOI: 10.1080/09638288.2018.1530307

To link to this article: https://doi.org/10.1080/09638288.2018.1530307

View supplementary material

Published online: 19 Feb 2019.

Submit your article to this journal

View Crossmark data
Validity and reliability of the Persian ABILHAND-Kids in a sample of Iranian children with cerebral palsy

Elham Mohammadkhani-Pordanjani, Carlyne Arnould, Parvin Raji, Noureddin Nakhostin Ansari and Scott Hasson

Aim: To develop a Persian version of ABILHAND-Kids and to determine its reliability and validity in Persian-speaking children with cerebral palsy (CP).

Method: The ABILHAND-Kids questionnaire was translated into Persian language and cross-culturally adapted following guidelines. The Persian ABILHAND-Kids was administered to 50 parents of CP children. Among the 50 parents of CP children, 30 of them participated in a test–retest reliability phase. Fifty parents of healthy children participated for discriminative validity.

Results: The Rasch analysis indicated the unidimensionality, reliability, and global invariance of the Persian ABILHAND-Kids. The internal consistency reliability was high (Cronbach's alpha = 0.96). Floor and ceiling effects were insignificant (4%). The Intraclass Correlation Coefficients of test–retest reliability were 0.96 and 0.70 for item difficulties and children's measures, respectively. The standard error of measurement and smallest detectable change for CP measure were 11.21 and 31.07%, respectively. The discriminative validity of the Persian ABILHAND-Kids was demonstrated by statistically significant lower ABILHAND-Kids measures in CP children than in healthy children (p < 0.001). Cross-cultural validity between the Persian and original version was established for 19 out of the 21 ABILHAND-Kids items.

Interpretation: The Persian ABILHAND-Kids questionnaire is reliable and valid for assessing manual ability in Persian speaking children with CP.

> IMPLICATIONS FOR REHABILITATION

- The Persian version of ABILHAND-Kids is developed and presented as a valid and reliable instrument for use by Persian-speaking clinicians and researchers.
- It is now possible for the Persian-speaking researchers to participate in international investigations and to compare Persian data with those from other countries.

Introduction

Cerebral palsy (CP) is a neurodevelopmental, nonprogressive condition that occurs in infancy and continues throughout the life of the affected children. CP is a common cause of physical disability in children and typically results in disorders of movement and posture [1]. The overall prevalence of CP worldwide has been reported as 2.11 per 1000 live births [2]. Patients with CP may have a wide range of dysfunctions, activity limitations, and reduced quality of life [3,4].

Arm and hand functions are impaired in about 60% of children with CP [5]. The integrity of the central nervous system (CNS) is essential for hand function and to properly perform daily tasks in various settings [6,7]. Individuals with CP having hand dysfunction experience serious difficulties in performing activities of daily living (ADL) [8,9].

There are various instruments for measuring hand functioning in children with CP (e.g. Manual Ability Classification System (MACS), Jebsen Taylor Test of Hand Function, the Assisting Hand Assessment (AHA), Melbourne Assessment of unilateral upper limb function, Quality of Upper Extremity Skills Test (QUEST), and ABILHAND-Kids questionnaire) [10,11]. The ABILHAND-Kids developed in 2004 is a reliable, valid, and responsive measure for assessing manual ability in children with upper limb impairments [5,12,13]. According to a recent systematic review about outcome measures evaluating hand function in children with CP, the strongest level of evidence for validity and reliability in the measurement of manual ability in daily activities was found for the ABILHAND-Kids questionnaire [14]. This questionnaire can be used in all levels of MACS [15], and is proper for goal setting in occupational therapy interventions [10].

The ABILHAND-Kids is a suitable and quick-to-administer scale which consists of 21 mainly bimanual items covering various...
domains of daily life. For each item, the parents are asked to provide their perceived child’s difficulty on a three-level scale: Impossible (score 0), Difficult (score 1), or Easy (score 2) [12].

The ABILHAND-Kids questionnaire has been validated in children with CP and is available in Belgian [12], French, Swedish, Polish (Test packages: www.rehab-scales.org/abilhand-kids-downloads), Dutch [16], English [17], Brazilian Portuguese [18], and Turkish [19]. The ABILHAND-Kids questionnaire is not available in the Persian Language. Therefore, the aim of the present study was to translate and adapt the ABILHAND-Kids into Persian language and to test the validity and reliability of a Persian ABILHAND-Kids in children with CP.

Methods

The review Board and the Ethics Committee of the School of Rehabilitation, Tehran University of Medical Science authorized the study protocol.

Translation and cross-cultural adaptation

The translation and cultural adaptation of ABILHAND-Kids was performed by the forward and backward translation procedure [20,21]. The original English ABILHAND-Kids was forward translated independently by two bilingual translators into Persian language. The research team and both translators discussed and agreed on a synthesized Persian version of the ABILHAND-Kids. Then, two other bilingual translators independently back-translated the agreed version of the Persian ABILHAND-Kids into English. The backward translators were blinded to the study and had no prior knowledge of the scale.

A pretest was conducted with the Persian ABILHAND-Kids for face and content validity in 30 parents of children with CP. The parents who participated in the pretest were not included in the validation phase of the study with the final Persian ABILHAND-Kids. The individuals reported no difficulties in the understanding of scale items, and consequently, the Persian ABILHAND-Kids was finalized for further evaluations of validity and reliability.

Study participants

The children with CP were recruited through the University Rehabilitation Centers in Tehran, Iran. To be included in the study, the children have to be diagnosed as CP, aged between 6 and 15 years old. Children with IQ <60 were excluded. Parents of children with CP were required to be able to read and write and gave their written informed consent. Table 1 presents demographic and clinical characteristics of the children with CP (n = 50).

Sample size

Fifty parents of CP children and 50 parents of healthy children were recruited in this study to satisfy the minimal sample size based on the quality criteria required for reliability and validity [22].

Procedure

An experienced occupational therapist interviewed the parents of children, recorded demographic data, and determined the children’s level of gross motor skills using the Gross Motor Function Classification System (GMFCS) [23]. The parents of eligible children with CP completed the Persian version of ABILHAND-Kids rating their child’s ease/difficulty in performing activities defined in the questionnaire. To investigate the scale’s discriminative validity, parents of healthy subjects also completed the Persian ABILHAND-Kids. Finally, 30 parents of children with CP (a subset from the original 50 parents) completed the Persian ABILHAND-Kids again one week later to study its test-retest reliability.

Statistical analysis

Rasch analysis

The psychometric qualities of the Persian ABILHAND-Kids was analyzed using the Rasch rating scale model through the RUMM2020 computer program (RUMM Laboratory Pty. Ltd, Perth, Western Australia) [24]. The probabilistic Rasch model [25], increasingly used in the development of measurement tools in the health field, estimates the capacity of each person and the difficulty of each item on a common linear scale from the answer of every person to every item [25]. The Rasch model verifies whether a scale meets the requirements of an objective measurement. It also converts the observed ordinal total scores into linear measures expressed in logits, a measurement unit that is constant and reproducible throughout the measurement scale. At any level of the measurement scale, a 1-logit difference in subjects’ ability implies a constant ratio of their odds of success ($e^{1.71}$) to any given item. To provide a more common and understandable measurement unit, the logit unit has been transformed into a centile scale where 0% represents the lowest manual ability level and 100% the highest manual ability level.

The criteria used to verify the psychometric qualities of the Persian ABILHAND-Kids were:

1. Relevant items: an item is considered as relevant when it belongs to the daily reality of most Persian children. Items with a response rate $\geq 80\%$ were considered as relevant [26].
2. Local independency: local dependency occurs when the scores attributed to the subjects to an item depend on the scores given to another item of the same scale [27]. Residual correlations between two items higher than 0.3 indicate local dependency. However, the local dependency is not a threat for the psychometric qualities of the scale if the absolute value of the average residual correlations is lower than 0.2 [28].
3. Unidimensionality: unidimensionality implies that the manual ability is the only attribute playing a role in choosing the response for an item. Items presenting standardized residuals values between $-2.5$ and $2.5$ or a $p$ values of the $\chi^2$ fit statistic (computed over 3 class intervals (CI) of increasing ability levels) $\geq 0.05$ were considered as unidimensional [24]. Due to

Table 1. Characteristics of children with cerebral palsy (n = 50).

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Mean ± SD (range)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (years)</td>
<td>7.9 ± 2.2 (6–15)</td>
</tr>
<tr>
<td>Sex</td>
<td></td>
</tr>
<tr>
<td>Boys</td>
<td>30</td>
</tr>
<tr>
<td>Girls</td>
<td>20</td>
</tr>
<tr>
<td>School education (years)</td>
<td>1.1 ± 1.7 (0–8)</td>
</tr>
<tr>
<td>Type of CP</td>
<td></td>
</tr>
<tr>
<td>Hemiplegia (left/right)</td>
<td>19 (12/7)</td>
</tr>
<tr>
<td>Diplegia</td>
<td>19</td>
</tr>
<tr>
<td>Quadriplegia</td>
<td>12</td>
</tr>
<tr>
<td>GMFCS</td>
<td></td>
</tr>
<tr>
<td>Level I: most independent in gross motor function</td>
<td>27</td>
</tr>
<tr>
<td>Level II</td>
<td>6</td>
</tr>
<tr>
<td>Level III</td>
<td>6</td>
</tr>
<tr>
<td>Level IV</td>
<td>4</td>
</tr>
<tr>
<td>Level V: least independent in gross motor function</td>
<td>7</td>
</tr>
</tbody>
</table>

SD: standard deviation; GMFCS: Gross Motor Function Classification System.
the number of statistical tests undertaken, Bonferroni corrections were applied to p values of fit statistics (i.e., p = 0.05/number of items = 0.0024) [29]. A Principal Component Analysis (PCA) was also performed on the residuals to identify the factor that best explains observed responses variations not attributable to manual ability. The scale is considered unidimensional if the first PCA residual factor accounts for less than 30% of the observed variance [30].

4. No differential item functioning (DIF): children with identical manual ability but different demographic or clinical characteristics should have the same probability of succeeding any particular item. If this is not the case, the item presents “differential functioning”. The invariance of the item difficulty hierarchy was tested for age (<8 years old versus ≥8 years old), gender, clinical form of CP (hemiplegia vs. diplegia vs. quadriplegia), and GMFCS level (GMFCS = I vs. GMFCS ≥ II) [23]. DIF was measured by computing a two-way analysis of variance (ANOVA) on the standardized residuals of 3 CIs of increasing ability levels; the first factor was the investigated patient characteristics (age, gender, clinical form of CP, GMFCS level) and the second factor was the CIs [31]. Items presenting a significant main effect for the first patient factor indicate a significant DIF. Due to the number of statistical tests undertaken, Bonferroni corrections were applied to p values of DIF statistics (i.e., p = 0.05/number of items = 0.0024) [29].

**Item-patient targeting and internal consistency reliability**

Item-patient targeting was verified by comparing mean children’s measures and mean item difficulty (arbitrary set at 0 logit or 50% of logits) to verify whether difficulty of the scale was globally well adapted to the sample. Item-patient targeting was also visually analyzed by comparing children’s measures and item thresholds’ distributions to detect potential gaps in the scale. Gaps would prevent the distinction of several children with different manual ability levels located in this area. A percentage of children with minimum (0% of logits) or maximum total score (100% of logits) ≥ 15% was considered as significant floor and ceiling effects [22].

The degree of precision achieved by the Persian ABILHAND-Kids (i.e., internal consistency reliability) was examined by computing both the traditional Cronbach’s alpha coefficient and the Rasch person separation index (PSI; range: 0–1). Cronbach’s alpha and PSI values ≥ 0.9 indicate high internal consistency. The PSI allows the number of manual ability levels that may be statistically distinguished in the sample to be calculated [32].

**Test–retest reliability**

The test–retest reliability was examined in a subsample of 30 parents of children with CP who answered a second time to the Persian ABILHAND-Kids questionnaire one week later. Children’s manual ability measures and the item difficulty hierarchy obtained at the first and second assessments were compared. The degree of agreement between children’s measures and items’ difficulties obtained at the first and the second assessments was computed using an intra-class coefficient (ICC, two-ways random effects model, absolute agreement, average measure) interpreted as fair (0.40–0.59), good (0.60–0.75), and excellent (≥ 0.75) [33]. To put the children’s measures and the items’ difficulties on the same scale for both assessments, the adjustment of the origin of both calibrations was obtained by anchoring the items/thresholds of the second calibration at the difficulty level of the first one [26].

The agreement measures of the standard error of measurement (SEM) and the smallest detectable change (SDC) were calculated using the formulas SD₁,1-ICC and 2.77 × SEM, respectively. The data were analyzed using the SPSS statistical software, version 17 (SPSS, Inc, Chicago, IL). The p < 0.05 was accepted as statistically significant.

**Discriminative validity**

The capacity of the Persian ABILHAND-Kids to discriminate children with CP from typically-developing (TD) children was investigated by comparing their manual ability measures using a Mann-Whitney test as normality and homoscedasticity were not satisfied (SPSS statistical software, version 17; level of significance set at p < 0.05). To put the children’s measures on the same scale, the adjustment of the origin was obtained by anchoring the items and the thresholds of the TD children’s calibration at the difficulty level of the CP children’s one.

**Cross-cultural validity**

Similar to the graduations of a metric ruler, ABILHAND-Kids items should have the same difficulty for all nationalities or cultures, regardless of the language version being used. A linguistic validation is not sufficient to prove cross-cultural validity as item difficulties may vary across countries due to translation errors or cultural differences. Cross-cultural validity is only obtained by verifying the absence of DIF between cultural groups. To investigate the DIF across countries (Iran vs. Belgium), the Persian data were pooled with the data of 50 Belgian children with CP randomly selected from the original data set including 113 children [12].

**Results**

**Sample characteristics**

Fifty children with CP (30 boys) with a mean age of 7.9 ± 2.2 years (range = 6–14 years) participated in this study. All parents who participated in this study were mothers. Table 1 shows the demographic and clinical characteristics of our sample of Persian children with CP.

**Description of Persian ABILHAND-Kids**

The definition and use of the Persian ABILHAND-Kids scale is depicted in Figure 1. The top panel shows the distribution of manual ability measures of the children as perceived by the parents. The bottom panel illustrates the sigmoidal (S-shaped) curve relationship between the finite total raw scores and the infinite manual ability measures (expressed in % and in logits). This relationship is approximately linear between total scores of 12 and 32. Outside this central range, however, a unitary progression in total score accounts for an increasing amount of manual ability measure. In the central range, the change in manual ability measure corresponding to a unitary increment in total score from 20 to 21 scores is equal to 1.36% (0.14 logits). Outside this central range, it increases to 8.33% (0.86 logits) for the same increment in total score from 0 to 1. This six-fold difference denotes the non-linearity of the total score. The conversion between raw total scores and linear measures of the Persian ABILHAND-Kids is given in Supplementary File. Note that the conversion table is only useful without missing data. The middle panel shows the expected response to a given item as a function of the underlying manual ability measure. For all items, the distance between the two thresholds was 17.23% (1.78 logits) as the rating scale model was used. By comparing the ability of a given child to the difficulty of each item, it is possible to determine the expected score of the child to the item. According to the parents’ perception, a child
with a total raw score of 20 will have an ability close to 50% (0 logit) and would be expected to perform without difficulty the five easiest activities, to perform with some difficulties the average activities, and to fail to perform the five most difficult activities.

**Metric properties of Persian ABILHAND-Kids**

The calibration of the 21-item Persian ABILHAND-Kids scale is presented in Table 2. The items are sorted, from top to bottom, in order of decreasing difficulty (range: 2.63 to −2.71 logits). Higher logit values indicate more difficult activities. Table 2 also reports the difficulty of the items expressed in centiles (% of logits), the standard error (SE) associated with each item difficulty expressed in logits (mean: 0.30 logits; range: 0.26–0.45 logits) or centiles (mean: 2.93%; range: 2.51–4.35%).

Two items (“Opening lid jar jam” and “Fastening snap jacket”) have more than 20% of missing values in our sample (22 and 26% of missing data, respectively).

There was no important local dependency (R > 0.70) between pairs of items. However, three small local dependencies were present between: “Buttoning up shirt” and “Buttoning up trousers” (R = 0.39); “Opening breadbox” and “Putting on hat” (R = 0.36); “Sharpening pencil” and “Opening bag chips” (R = 0.32). However, the absolute value of the average residual correlations was lower than 0.2 (absolute value of −0.049) indicating that this small local dependency is not a threat for the psychometric qualities of the scale.
Table 2. Persian ABILHAND-Kids calibration for children with cerebral palsy.

<table>
<thead>
<tr>
<th>Item</th>
<th>Difficulty (logits)</th>
<th>Difficulty (% of logits)</th>
<th>SE (logits)</th>
<th>SE (% of logits)</th>
<th>Standardized residuals</th>
<th>Chi-square ($\chi^2$)</th>
<th>Probability</th>
</tr>
</thead>
<tbody>
<tr>
<td>Buttoning up trousers</td>
<td>2.63</td>
<td>76.11</td>
<td>0.30</td>
<td>2.89</td>
<td>-0.55</td>
<td>1.65</td>
<td>0.44</td>
</tr>
<tr>
<td>Buttoning up shirt</td>
<td>1.94</td>
<td>69.42</td>
<td>0.28</td>
<td>2.68</td>
<td>-1.34</td>
<td>3.23</td>
<td>0.20</td>
</tr>
<tr>
<td>Rolling up sleeve</td>
<td>1.39</td>
<td>64.05</td>
<td>0.27</td>
<td>2.57</td>
<td>0.38</td>
<td>0.11</td>
<td>0.95</td>
</tr>
<tr>
<td>Zipping up jacket</td>
<td>0.95</td>
<td>59.79</td>
<td>0.26</td>
<td>2.55</td>
<td>0.81</td>
<td>4.98</td>
<td>0.08</td>
</tr>
<tr>
<td>Fastening snap jacket</td>
<td>0.92</td>
<td>59.56</td>
<td>0.30</td>
<td>2.89</td>
<td>0.78</td>
<td>2.25</td>
<td>0.32</td>
</tr>
<tr>
<td>Zipping up trousers</td>
<td>0.73</td>
<td>57.66</td>
<td>0.27</td>
<td>2.65</td>
<td>-0.79</td>
<td>2.72</td>
<td>0.26</td>
</tr>
<tr>
<td>Putting backpack on back</td>
<td>0.64</td>
<td>56.82</td>
<td>0.26</td>
<td>2.54</td>
<td>1.60</td>
<td>6.54</td>
<td>0.04</td>
</tr>
<tr>
<td>Opening cap chips</td>
<td>0.52</td>
<td>55.68</td>
<td>0.26</td>
<td>2.51</td>
<td>-0.81</td>
<td>2.90</td>
<td>0.23</td>
</tr>
<tr>
<td>Opening lid jar jam</td>
<td>0.51</td>
<td>55.61</td>
<td>0.30</td>
<td>2.88</td>
<td>0.62</td>
<td>10.29</td>
<td>0.01</td>
</tr>
<tr>
<td>Putting toothpaste on toothbrush</td>
<td>0.48</td>
<td>55.30</td>
<td>0.28</td>
<td>2.68</td>
<td>-0.15</td>
<td>1.08</td>
<td>0.58</td>
</tr>
<tr>
<td>Washing upper body</td>
<td>0.39</td>
<td>54.33</td>
<td>0.26</td>
<td>2.53</td>
<td>0.35</td>
<td>0.62</td>
<td>0.73</td>
</tr>
<tr>
<td>Sharpening pencil</td>
<td>0.37</td>
<td>54.16</td>
<td>0.28</td>
<td>2.68</td>
<td>-1.06</td>
<td>4.35</td>
<td>0.11</td>
</tr>
<tr>
<td>Taking off T-shirt</td>
<td>-0.10</td>
<td>49.67</td>
<td>0.27</td>
<td>2.61</td>
<td>0.50</td>
<td>1.99</td>
<td>0.37</td>
</tr>
<tr>
<td>Unwrapping chocolate bar</td>
<td>-0.36</td>
<td>47.14</td>
<td>0.27</td>
<td>2.65</td>
<td>0.02</td>
<td>0.57</td>
<td>0.75</td>
</tr>
<tr>
<td>Opening cap toothpaste tube</td>
<td>-0.69</td>
<td>43.94</td>
<td>0.29</td>
<td>2.84</td>
<td>-1.05</td>
<td>0.24</td>
<td>0.89</td>
</tr>
<tr>
<td>Unscrewing cap bottle</td>
<td>-0.70</td>
<td>43.83</td>
<td>0.29</td>
<td>2.85</td>
<td>0.62</td>
<td>8.81</td>
<td>0.05</td>
</tr>
<tr>
<td>Pouring water</td>
<td>-1.25</td>
<td>38.50</td>
<td>0.33</td>
<td>3.19</td>
<td>-0.06</td>
<td>0.63</td>
<td>0.73</td>
</tr>
<tr>
<td>Coin out pocket</td>
<td>-1.34</td>
<td>37.62</td>
<td>0.33</td>
<td>3.23</td>
<td>-0.28</td>
<td>3.95</td>
<td>0.14</td>
</tr>
<tr>
<td>Putting on hat</td>
<td>-2.10</td>
<td>30.31</td>
<td>0.37</td>
<td>3.60</td>
<td>-1.14</td>
<td>1.14</td>
<td>0.56</td>
</tr>
<tr>
<td>Switching on bedside lamp</td>
<td>-2.20</td>
<td>29.29</td>
<td>0.42</td>
<td>4.10</td>
<td>0.24</td>
<td>1.44</td>
<td>0.49</td>
</tr>
<tr>
<td>Opening breadbox</td>
<td>-2.71</td>
<td>24.44</td>
<td>0.45</td>
<td>4.35</td>
<td>-0.35</td>
<td>1.35</td>
<td>0.51</td>
</tr>
</tbody>
</table>

No major misfit was found in the items as indicated by the goodness-of-fit statistic tests (all standardized residuals values contained between −2.5 and 2.5; all $\chi^2$ p values > 0.0024; Table 2). Moreover, the scale presented a mean $\chi^2$ probability of 0.40 indicating that overall, the 21 items contribute to the definition of a unidimensional measure of manual ability. This is confirmed by the PCA on the residuals showing than the first residual factor (not attributable to manual ability) accounts for only 13% of the observed variance.

According to the Bonferroni adjustment, no major DIF was present indicating that the scale is used in the same way for different demographic (age, gender) and clinical (clinical form of CP, GMFCS level) subgroups of individuals as the item difficulty hierarchy is invariant.

Item-patient targeting and internal consistency reliability

The mean item difficulty was equal to 50 ± 13.26% (0 ± 1.37 logits) and the mean children’s manual ability measure was equal to 56 ± 19.94% (0.54 ± 2.06 logits) meaning that the difficulty of the scale was well adapted to our sample. Top and upper middle panels of Figure 1 show the good item-patient targeting. No major gaps were present on the Persian ABILHAND-Kids manual ability measurement scale from 16 to 85% (from −3.60 to 3.52 logits). According to the distribution of subject measures, 30% of the children in our sample should successfully perform all the listed activities easily or with some difficulty. Six percent of the children should perform all activities easily and 6% should not be able to perform any of the 21 Persian ABILHAND-Kids items (Figure 1). Therefore, the range of difficulties of the ABILHAND-Kids items fits the distribution of children’s abilities. Only two parents of Persian children with CP (4%) reported a minimum total score (0% of logits) on the questionnaire. The manual ability of their children cannot be measured by the Persian ABILHAND-Kids scale because all activities are impossible. No parent reported maximal total score (100% of logits) on the questionnaire. Consequently, no significant floor and ceiling effects were observed on the Persian ABILHAND-Kids.

The Cronbach’s alpha coefficient for Persian ABILHAND-Kids was high (0.963). The Cronbach’s alpha values if an item was deleted were ranged between 0.960 and 0.963. The PSI was equal to 0.946 indicating that the scale has a good precision enabling approximately 6 manual ability levels to be statistically distinguished in our sample.

Test-Retest reliability

Figure 2 illustrates test-retest of items’ and persons’ locations based on the 30 parents’ responses. The items’ difficulties at first assessment were highly correlated with difficulties at second assessment (ICCagreement: 0.96; CI 95%: 0.90–0.98; $p < 0.001$). All items lie within the 95% of confidence interval indicating that they have the same estimated difficulties at the first and second assessments. The children’s measures at first assessment presented also good agreement with the children’s measures at second assessment (ICCagreement: 0.7; CI 95%: 0.33–0.85; $p < 0.001$). The SEM and SDC for CP measure were equal to 11.21% (1.16 logits, raw score of 2.21) and 31.07% (3.21 logits, raw score of 6.13), respectively.

Discriminative validity

Discriminative validity was assessed by comparing the manual ability measures of 50 parents of children with CP and the manual ability measures of 50 TD subjects (26 girls; age (mean ± SD range): 8.4 ± 2.2 [6–14] years). A significant difference between CP (median [Q1–Q3]: 61.33% [46.19–67.79%], range = 0–91.50%) and TD manual ability measures (median [Q1–Q3]: 100% [81.80–100%], range = 67.80–100%) were found ($p < 0.001$) indicating that the Persian ABILHAND-Kids is able to discriminate children with CP from TD children.

Cross-cultural validity

The invariance of the item difficulties was investigated across countries (Iran vs. Belgium) by pooling the Persian data with the data of 50 Belgian children with CP. All items have less than 15% of missing values. There was no important local dependency ($R > 0.70$) between pairs of items. However, one small local dependency was present between: “Buttoning up shirt” and “Buttoning up trousers” ($R = 0.38$). However, the absolute value of the average residual correlations was lower than 0.2 (absolute value of −0.048) indicating that this small local dependency is not a threat for the psychometric qualities of the scale. No major
mismatch was found in the items as indicated by the fit statistics (all standardized residuals values contained between $-2.5$ and $2.5$; all $\chi^2$ $p$ values $>0.0024$). All 21 items contribute to the definition of a unidimensional measure of manual ability. This is confirmed by the PCA on the residuals showing than the first residual factor (not attributable to manual ability) accounts for only 12% of the observed variance. The mean item difficulty was equal to $50 \pm 13.07\%$ ($0 \pm 1.35$ logits) and the mean children of the observed variance. The mean item difficulty was equal to $59 \pm 23.43\%$ ($0.89 \pm 2.42$ logits) meaning that the difficulty of the scale was well adapted to the pooled sample. Five parents of children with CP (2% of Persians and 3% of Belgians) reported a minimum total score (0% of logits) on the questionnaire and four parents (4% of Belgians) given a maximal total score (100% of logits). The PSI was equal to 0.961 indicating that the scale has a good precision enabling approximately 7 manual ability levels to be statistically distinguished in the pooled sample. According to the Bonferroni adjustment, no major DIF was present across age, gender, clinical form of CP, and GMFCS level.

However, two major DIFs were observed across countries: “Unwrapping chocolate bar” ($p=0.0004$) which was easier for Belgians and “Opening bread box” ($p=0.0001$) which was easier for Iranians. Figure 3 shows the item characteristic curves (solid lines representing the score expected by the Rasch model according to the manual ability) and mean observed scores (dots) for 3 interval class of CP children for Belgian (gray circles) and Iranian (white triangles) children. Panels A and B illustrate two items (“Sharpening pencil” and “Buttoning up shirt”, respectively) presenting no DIF. The mean scores observed for the children (in the 3 increasing ability levels) of the two countries present similar patterns that superimposed meaning that the difficulty of these items does not vary according to the children’s culture. Panels C and D illustrate the two items (“Unwrapping chocolate bar” and “Opening bread box”, respectively) presenting a major DIF across countries. The mean scores observed for the children of both countries evolve in parallel. For a given manual ability level, the score observed on “Unwrapping chocolate bar” (panel C) is systematically higher for Belgians and systematically lower for Iranians. In other words, this item is systematically easier for Begians and more difficult for Iranians. It is the opposite for “Opening bread box” (panel D). These results indicate that the difficulty of these two items changes from one country to the other highlighting the importance to use the Persian calibration in Iran.

**Discussion**

The present study was sought to perform the cultural adaptation of the ABILHAND-Kids questionnaire into Persian language and to explore its reliability and validity. The priori hypotheses regarding the reliability and validity of the new Persian ABILHAND-Kids were verified. Hence, the Persian ABILHAND-Kids instrument which satisfy the requirements of an objective measurement showed satisfactory reliability and validity like the original and translated versions of ABILHAND-Kids [18,19,26]. Our findings suggest that the Persian ABILHAND-Kids questionnaire can be used to assess the Persian-speaking children with CP with manual performance disturbances.

In this study, the translation and cross-cultural adaptation process was followed to keep the same original concepts in Persian culture. The prefinal version of the Persian ABILHAND-Kids did not require any alterations and all items were answered, emphasizing the content validity of the questionnaire. The acceptability of the Persian ABILHAND-Kids was confirmed in the final version of the questionnaire as subjects answered all the questions. Two items (“Opening lid jar jam” and “Fastening snap jacket”) had slightly more than 20% of missing values in our sample (22 and 26% of missing data, respectively) showing that these items might be slightly less relevant in Persian sample than in Belgian sample. In Iran, it is the mothers who usually open the lid of a jar of jam. As well, Iranians wear rarely clothes with snap fasteners but rather clothes with zipper or buttons. So, cultural differences may explain the higher rate of missing values observed in some items of the Persian ABILHAND-Kids compared to the Belgian version. However, the percentage of missing values remains reasonable (≤26%).

The 21 items of the Persian ABILHAND-Kids provide a unidimensional measure of manual ability expressed on an interval-level scale (in logits or in %) enabling the manual ability of different subjects to be quantitatively compared or manual ability changes to be monitored over time. Such as the original ABILHAND-Kids scale, the item difficulty hierarchy of the Persian
version is congruent with clinical observations strengthening the validity of the scale.

Although no important local dependency was observed between pairs of items, three small local dependencies were observed ($R \leq 0.39$). “Buttoning up shirt” and “Buttoning up trousers” require the same kind of movements to be achieved; “Opening breadbox” and “Putting on hat” are bimanual activities that can be managed in several unimanual steps involving a certain amount of gross manual ability; “Sharpening pencil” and “Opening bag chips” are two activities requiring a digital activity from both hands. These small local dependencies do not affect the psychometric qualities of the scale as shown by the insignificant mean of residual correlations. This is consistent with the absence of local dependency observed in the Turkish version of ABILHAND-Kids [12,19].

The scale is also invariant between groups allowing the measurement of manual ability in children with CP whatever their age, gender, clinical form of CP, and GMFCS level. Moreover, Persian ABILHAND-Kids presents an excellent degree of precision since the 21 items are well targeted on our sample expanding a wide range of functional states (Cronbach’s alpha: 0.96; PSI: 0.95). The Persian ABILHAND-Kids has the potential to statistically distinguish approximately 6 different manual ability levels in children with CP. Similar results were found for the original Belgian (PSI: 0.94) and Brazilian-Portuguese (Cronbach’s Alpha: 0.99) ABILHAND-Kids versions [12,18,19]. Only the Turkish version of ABILHAND-Kids [17] showed lower precision (Cronbach’s alpha = 0.81). Consistently with other versions of ABILHAND-Kids [19,26], no significant floor and ceiling effects were observed on the Persian ABILHAND-Kids. The lack of floor and ceiling effects indicate the potential of the Persian ABILHAND-Kids to detect any sensible change in manual ability induced, for example, by surgery, rehabilitation, biomedical treatment, or the use of assisting devices [34]. According to their parents, only 4% of the children with CP were unable to perform at least one item. These two children were tetraplegic/paretic and were severely limited even with the use of an assistive technology (GMFCS: Level V).

Good test–retest reliability was also observed for Persian ABILHAND-Kids. This is in accordance with the original Belgian ($r = 0.91$) and the Turkish (ICC = 0.94) ABILHAND-Kids versions [12,19]. However, the test–retest reliability originally reported in the Belgian version was based on a correlation coefficient that does not consider the systematic errors between measurements and source of errors [12]. The good test–retest reliability found for the Persian ABILHAND-Kids indicates that when the clinicians administer this tool in stable subjects on two or more occasions, similar scores with some variability would be obtained [35]. Test–retest reliability is an important quality for an instrument as the clinicians must be assured that the changes observed after an intervention are due to the intervention and not to random variations from the instrument. In the present study, the SEM and SDC values in terms of raw scores were acceptable and relatively similar to those of Turkish ABILHAND-Kids (SEM = raw score of 2.6; SDC = raw score of 7.1) [19]. The SEM indicates the degree of errors in an individual score on the Persian ABILHAND-Kids. In this study, the amount of SEM was smaller than the Turkish version of ABILHAND-Kids indicating a lesser variability of scores in repeated measurements and subsequently, a higher reliability of the Persian ABILHAND-Kids. The SEM was used to calculate the SDC which reflects the smallest real within-individual change in the Persian ABILHAND-Kids measure/score that is above the measurement error. Hence, a change of more than six points on the Persian ABILHAND-Kids indicates a true change, exceeding the error variance within the measure. Once again, the SDC found for the Persian ABILHAND-Kids was smaller than the one found in the Turkish version [19]. The SEM and SDC were not reported for the original Belgian and other versions of ABILHAND-Kids [12,18,23,25]. Though acceptable, the SDC for the Persian ABILHAND-Kids is somewhat large. Indeed, a change of more than 31% of logits (slightly less than one third of the measurement range) on the Persian ABILHAND-Kids is required to be considered as a true change meaning that the scale might not be sensitive enough. This seems surprising considering the excellent internal consistency of the scale that should give it the potential

Figure 3. Item characteristic curves (solid lines representing the score expected by the Rasch model according to the manual ability level) and mean observed scores (dots) for 3 interval classes of people (Belgian children represented by gray circles and Persian children by white triangles) with increasing manual ability levels allowing to investigate differential item functioning (DIF) across countries. Panels A and B: item characteristic curves of items “Sharpening pencil” (Panel A) and “Buttoning up shirt” (Panel B) which present no DIF across countries. Panels C and D: item characteristic curves of items “Unwrapping chocolate bar” (panel C) and “Opening breadbox” (Panel D) which present a DIF across countries.
to detect small clinical changes. In our sample, a small proportion of parents failed to assess consistently their child (2 parents and, to a lesser extent, 3 additional parents). The inconsistent responses of these marginal parents might negatively influence the calculation of the SDC. So, a further study with a larger sample size is needed to confirm the SDC of the Persian ABILHAND-Kids scale.

The discriminative validity of the Persian ABILHAND-Kids was proved as the scale was able to discriminate children with CP from TD children ($p < 0.001$). The discriminative validity of ABILHAND-Kids was not evaluated in the original and other adapted versions of the scale [5,12,18,19].

This study is the first one that investigates the cross-cultural validity of ABILHAND-Kids through Rasch analysis. Indeed, Rasch analysis was not performed for the Brazilian Portuguese ABILHAND-Kids [18], and cross-cultural analysis was not reported for the Turkish ABILHAND-Kids [19]. In our study, two major DIFs across countries were observed indicating that the difficulty of two items changes from one country to the other. “Unwrapping chocolate bar” was systematically easier for Belgians and “Opening bread box” was systematically easier for Iranians. These findings might be explained by traditional and lifestyle differences between both countries. Although the item difficulty hierarchy is globally similar between the original [12] and the Persian ABILHAND-Kids versions strengthening the construct validity of the scale, some slight discrepancies are also observed. For instance, “Opening lid jar jam” and “Putting on hat” are slightly easier for Iranians. This highlights the importance to use the Persian calibration in Iran.

The sample size ($n = 50$) used to calibrate the Persian ABILHAND-Kids is small and may potentially lead to imprecise estimates. Nevertheless, the standard errors associated with the items difficulties (mean: 0.3 logits, 2.93%) is, on average, small indicating that the number of observations are sufficient to give useful item calibrations. Larger sample size is also generally required to perform DIFs analyses.

Conclusion

The Persian ABILHAND-Kids is a highly reliable and valid tool for assessing manual ability in children with CP. The Persian ABILHAND-Kids can thus be used as an outcome measure in the clinics and research in Persian-speaking countries if the intervention is targeting manual ability in children with CP. However, further analyses including a larger number of observations are wanted to confirm our results and responsiveness of the Persian ABILHAND-Kids questionnaire should also be considered in the future.

Acknowledgements

The authors are grateful to the parents for participating in the study. We also thank the Research Deputy, Tehran University of Medical Sciences.

Disclosure statement

No potential conflict of interest was reported by the authors.

References


Terwee CB, Bot SD, de Boer MR, et al. Quality criteria were proposed for measurement properties of health status questionnaires. J Clin Epidemiol. 2007;60:34–42.


