Effective use of the assisting hand in adolescents with cerebral palsy

Louwers, A.M.

Citation for published version (APA):
Louwers, A. M. (2018). Effective use of the assisting hand in adolescents with cerebral palsy

General rights
It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations
If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: http://uba.uva.nl/en/contact, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.
Reliability of the Assisting Hand Assessment in adolescents

Annoek louwers
Lena Krumlinde-Sundholm
Karin Boeschoten
Anita Beelen

Developmental Medicine & Child Neurology 2017
Abstract

Aim
To investigate the interrater and test–retest reliability of the Assisting Hand Assessment in adolescents (Ad-AHA) with cerebral palsy (CP) and to evaluate the alternate-form reliability of different test activities.

Method
Participants were 112 adolescents with unilateral CP (60 males, 52 females; mean age 14y 5mo [standard deviation {SD} 2y 8mo], Manual Ability Classification System levels I–III). Reliability was evaluated using intraclass correlation coefficients (ICC), smallest detectable change (SDC), and Bland–Altman plots.

Results
ICCs for interrater (n=38) and test–retest reliability (n=31) were excellent: 0.97 (95%CI 0.94–0.98) and 0.99 (95% CI 0.98–0.99) respectively. The alternate-form reliability of different test activities was excellent for children (age 10–12y, n=30) performing the School-Kids AHA and Ad-AHA Board Game 0.99 (95% CI 0.98–0.99) and for adolescents (age 13–18y) performing the Ad-AHA Board Game compared to the Ad-AHA Present (n=28) 0.99 (95% CI 0.95–0.98), or the Ad-AHA Sandwich (n=29) 0.99 (95% CI 0.98–0.99) tasks. SDC for test–retest was 4.5 AHA-units.

Interpretation
Ad-AHA scores are consistent across different raters and occasions. The good alternate-form reliability indicates that the different test activities can be used interchangeably in adolescents with unilateral CP. Differences greater than or equal to 5 AHA-units can be considered a change beyond measurement error. The use of logit based AHA-units makes change comparable for persons at different ability levels.
Introduction

The upper limb dysfunction of children and adolescents with unilateral cerebral palsy (CP) may make the performance of activities involving reaching, grasping, and manipulating objects difficult. A common goal of treatment for adolescents with unilateral CP is to improve bimanual performance, because the effective use of the affected arm and hand in combination with the well-functioning hand is important for independence. A valid and reliable tool to assess arm-hand performance is needed in order to plan interventions, determine the effectiveness of different interventions, and monitor the development of hand-use of individuals with unilateral CP.

Within the group of individuals with unilateral CP, assessments of hand-use have primarily been developed for younger children with unilateral CP. Currently only the Melbourne Assessment 2 (MA2), Shriners Hospital for Children Upper Extremity Evaluation (SHUEE), and ABILHAND-Kids can be used in adolescents. But the MA2 and SHUEE mainly assess capacity, and the person-centred questionnaire ABILHAND-Kids assesses perceived performance difficulties, which can be quite different from the observed performance.1–3

Recently, an adjusted and improved version of the Assisting Hand Assessment (AHA), the Kids-AHA 5.0, demonstrated strong evidence of internal construct validity, 4 and the same scoring criteria can also be used to assess the bimanual performance of adolescents in a newly developed test situation, the board game ‘Go with the Floe’ (Ad-AHA Board Game).5 The AHA is commonly used to assess the quality of use of the affected hand during bimanual performance in individuals with unilateral CP. The Kids-AHA has been validated for children with unilateral CP aged 1 year 6 months to 12 years and has proven reliable and sensitive enough to monitor changes over time.6–11 With the newly developed test situation of the Ad-AHA, there is strong evidence that the AHA 5.0 is a valid instrument for assessing the bimanual performance of individuals aged 18 months to 18 years (AHA 18–18).5 The AHA is used to score object-related hand actions observed during a semi-structured test situation (10–20min) that elicits the use of both hands and which is suitable for different age groups. The AHA 18–18 includes different test activities: (1) with the ‘Small-Kids AHA’, children aged 18 months to 5 years engage in exploratory play with toys; (2) with the ‘School-Kids AHA’, children aged 6 to 12 years handle the same objects but within the context of two AHA adventure board games; and (3) with the Ad-AHA Board Game, adolescents aged 13 to 18 years handle age-appropriate objects while playing the board game ‘Go with the Floe’. Furthermore, two test activities have been developed in an adult post-stroke version of the AHA (Ad-AHA Stroke) with unpublished evidence of validity (L. Krumlinde-Sundholm, B. Lindquist, J. Plantin & B. Hoare, manuscript in preparation): (1) (un-) wrapping a present (Ad-AHA Present) and (2) preparing a sandwich (Ad-AHA Sandwich), which may also be suitable for adolescents.

The Small-Kids AHA and the School-Kids AHA have excellent interrater, intrarater, test–retest, and alternate form reliability.7,8 However, for the Ad-AHA Board Game, there is currently no evidence regarding inter- and intrarater agreement, test–retest stability, or interchangeability of test activities. This information is essential in order to be able to evaluate whether changes after treatment or follow up are real effects or the result of measurement error.

The aims of this study were to investigate different aspects of the reliability of the Ad-AHA for adolescents with unilateral CP concerning: (1) interrater, test–retest reliability and the smallest detectable change; (2) alternate-form reliability of the Ad-AHA Board Game.
versus School-Kids AHA (age group 10–13y); (3) alternate-form reliability of the Ad-AHA Board Game versus Ad-AHA Present and the Ad-AHA Board Game versus Ad-AHA Sandwich (age group 13–18y).

Methods

Participants

A convenience sample of 112 children and adolescents with unilateral CP, aged 10 to 18 years, were recruited by therapists working in eight different rehabilitation centers in Australia, Sweden, and the Netherlands (see Acknowledgements at end of paper) between May 2012 and February 2016. Forty-two of the children and adolescents participated in more than one type of reliability evaluation. The demographic characteristics and distribution of the adolescents are presented in Table I.

All participants and their primary caregivers gave informed consent. Participation was voluntary, and the parents and adolescents were informed that they could withdraw from the study at any stage. The ethical review committees of the involved centers approved the study.

Instrumentation

The AHA 5.0 was used, which consists of 20 items that describe object-related hand actions scored on a four-point rating scale. The same AHA scoring criteria can be used for both children and adolescents, but scored from different age-appropriate test activities, thereby covering the age range 18 months to 18 years (AHA 18–18). The sum scores are converted to AHA-units (interval logit measures) ranging from 0 to 100.

Procedure

All AHA assessments were videotaped according to the standard protocol described in the AHA manual and sent to one location for central scoring. The participants were tested once or twice depending on the type of reliability evaluated. None of the participants received any form of intensive therapy between the two test occasions. The examiners were blinded to the results of each other's and their own previous assessments. All videos were scored in random order.

To evaluate interrater reliability, the first 38 participants in the convenience sample played the Ad-AHA Board Game once and two raters independently scored the videotaped session. Both raters (occupational therapists) had at least 10 years of clinical experience and were trained in scoring the AHA 5.0.

To evaluate test–retest reliability, the adolescents played the Ad-AHA Board Game twice conducted by the same assessor, in the same environment with an interval of 1 to 2 weeks, and the videos were scored by the same assessor who was blind to previous obtained AHA scores. The test-retest condition thus includes stability of both the examined person’s behavior and intrarater agreement. For that reason, a separate intrarater analysis was not performed.

The four different test activities (School-Kids AHA, Ad-AHA Board Game, Ad-AHA Present, and Ad-AHA Sandwich) elicit bimanual actions involving the same underlying trait.
Alternate-form reliability determines whether the scores for the outcome measures of these tests are consistent relative to each other. In this study, agreement between AHA scores was analysed by comparing scores on the Ad-AHA Board Game with scores on one of the other three test activities. Two out of 85 adolescents performed three different test activities (School-Kids AHA, Ad-AHA Board Game, and Ad-AHA Present). The maximum interval between two tests was 2 weeks.

Statistical analyses

Analyses were performed with SPSS software, version 21.0 (IBM Corp., Armonk, NY, USA).

Agreement between raters and stability across time were calculated using the intraclass correlation coefficient (ICC; type 2.1, a two-way random effects single measures model of absolute agreement) with associated 95% confidence intervals (CIs). Agreement between alternate test forms was calculated using ICC type 3.1 (a two-way mixed effects single measures model of absolute agreement). Commonly accepted minimal standards for reliability coefficients are 0.70 for group comparisons and 0.90 to 0.95 for making clinical decisions for individuals. Portney and Watkins suggested that for many clinical measurements reliability should exceed 0.90 to ensure reasonable validity. In this study, ICC values higher than 0.90 were considered excellent, between 0.75 and 0.90 good, between 0.60 and 0.75 moderate, and lower than 0.60 poor.

The standard error of the measurement (SEM) is the amount of error that can be considered random measurement error. The SEM agreement was calculated using variance components from the ANOVA analysis as the square root of the pooled MeanSquare-time and MeanSquare-person 9 time (SEM = \sqrt{s^2_{t} + s^2_{residual}}). The smallest detectable change (SDC) indicates a change beyond measurement error. The SDC was calculated from the SEM of test–retest and alternate-form scores, using AHA-units (SDC = SEM x 1.96 x √2). At an individual level, a difference equal to or greater than the SDC can be considered a real change.

The presence of heteroscedasticity (or nonuniformity of error) was visually inspected with Bland–Altman plots.

In accordance with recommendations, the outcome of the AHA is reported in the Rasch-derived logit-based 0 to 100 AHA-unit scale (AHA-units). The ICC for the total measure, the SEM, and the SDC are given in the same units as the outcome measure (i.e. AHA-units 0–100 scale). To explore the agreement on item level, item’s raw scores were used.

Results

Interrater and test–retest reliability of the Ad-AHA Board Game

The ICC2.1 for interrater reliability was 0.97 (95% CI 0.94–0.98), indicating excellent agreement for the total AHA scores of 38 adolescents. With a mean interval of 9 days (range 7–14d), the test–retest ICC2.1 was also excellent, 0.99 (95% CI 0.98–0.99) (Table II).

The test–retest ICCs of all items (using raw scores) were good to excellent. Between two raters only 3 out of 20 items had moderate ICCs (Table II).
Table I. Demographic characteristics and distribution of participants to different reliability trials

<table>
<thead>
<tr>
<th>Type of reliability</th>
<th>n</th>
<th>Mean age (SD)</th>
<th>Sex Male (n)</th>
<th>Affected side, right (n)</th>
<th>MACS level (n)</th>
<th>GMFCS level (n)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>age years and months (SD)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Inter-rater reliability</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ad-AHA Board game</td>
<td>38</td>
<td>13.1 (2.8)</td>
<td>19</td>
<td>14</td>
<td>5</td>
<td>19</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>7</td>
<td>7</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>25</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0</td>
<td>7</td>
</tr>
<tr>
<td><strong>Test-retest reliability</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ad-AHA Board game</td>
<td>31</td>
<td>14.6 (2.2)</td>
<td>15</td>
<td>17</td>
<td>7</td>
<td>15</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>21</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0</td>
<td>5</td>
</tr>
<tr>
<td><strong>Alternate-form reliability</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ad-AHA Board game vs. School-Kids AHA</td>
<td>30</td>
<td>11.8 (1.4)</td>
<td>22</td>
<td>17</td>
<td>6</td>
<td>14</td>
</tr>
<tr>
<td>Ad-AHA Board game vs. Ad-AHA Sandwich</td>
<td>29</td>
<td>16.1 (2.0)</td>
<td>16</td>
<td>18</td>
<td>5</td>
<td>20</td>
</tr>
<tr>
<td>Ad-AHA Board game vs. Ad-AHA Present</td>
<td>28</td>
<td>15.1 (2.2)</td>
<td>9</td>
<td>16</td>
<td>8</td>
<td>15</td>
</tr>
</tbody>
</table>

MACS, Manual Ability Classification System; GMFCS, Gross Motor Function Classification System; N/a, missing data; Ad-AHA, Assisting Hand Assessment in adolescents.

Alternate-form reliability

For all alternate-form conditions the ICCs for individual AHA items were good to excellent (Table II). The ICC3.1 for alternate-form reliability of the School-Kids AHA versus the Ad-AHA Board Game was 0.99 (95% CI 0.98-0.99). It was 0.99 (95% CI 0.95-0.99) for the Ad-AHA Board Game versus Ad-AHA Present, and 0.99 (95% CI 0.98–0.99) for the Ad-AHA Board Game versus Ad-AHA Sandwich (Table II). Visual inspection of the Bland–Altman plots (Fig. 1) showed no heteroscedasticity.

Measurement error

The SEM for interrater and test–retest trials was 2.3 and 1.6 AHA-units respectively. The SEM for test–retest for the Ad-AHA Board Game resulted in a SDC of 4.5 AHA-units. The SEM and SDC for the four alternate forms are presented in Table II. At an individual level, a difference in scores greater than or equal to 5 AHA-units can be regarded as being change beyond measurement error (i.e., exceeding the random variation caused by intraperson differences in behaviour and by rater inconsistency) when using one of the four test activities (Ad-AHA Board Game, School-Kids AHA, Ad-AHA Sandwich, or Ad-AHA Present) (Table II).
Figure I. Bland–Altman plots: difference against mean for Assisting Hand Assessment (AHA) measures (in AHA-units)

Solid line: group mean difference. Dotted line: limits of agreement.
### Table II. Agreement between raters and alternate test forms, stability across time and smallest detectable change

<table>
<thead>
<tr>
<th>Measurement agreement</th>
<th>Interrater Ad-AHA Board Game</th>
<th>Test-retest Ad-AHA Board Game</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n=38</td>
<td>n=31</td>
</tr>
<tr>
<td><strong>Initiates use</strong></td>
<td>.80 (.64 - .89)</td>
<td>.94 (.88 - .97)</td>
</tr>
<tr>
<td><strong>Amount of use</strong></td>
<td>.63 (.39 - .78)</td>
<td>.79 (.61 - .89)</td>
</tr>
<tr>
<td><strong>Chooses AH when closer to objects</strong></td>
<td>.84 (.71 - .91)</td>
<td>.86 (.73 - .93)</td>
</tr>
<tr>
<td><strong>Stabilizes by weight or support</strong></td>
<td>.75 (.55 - .87)</td>
<td>.81 (.63 - .90)</td>
</tr>
<tr>
<td><strong>Reaches</strong></td>
<td>.84 (.70 - .92)</td>
<td>.93 (.85 - .96)</td>
</tr>
<tr>
<td><strong>Moves upper arm</strong></td>
<td>.66 (.43 - .81)</td>
<td>.92 (.85 - .96)</td>
</tr>
<tr>
<td><strong>Moves forearm</strong></td>
<td>.74 (.56 - .86)</td>
<td>.97 (.94 - .99)</td>
</tr>
<tr>
<td><strong>Grasps</strong></td>
<td>.80 (.65 - .89)</td>
<td>.95 (.90 - .98)</td>
</tr>
<tr>
<td><strong>Holds</strong></td>
<td>.86 (.75 - .93)</td>
<td>.94 (.87 - .97)</td>
</tr>
<tr>
<td><strong>Stabilizes by grasp</strong></td>
<td>.84 (.65 - .92)</td>
<td>.97 (.93 - .98)</td>
</tr>
<tr>
<td><strong>Readjust grasp</strong></td>
<td>.96 (.92 - .98)</td>
<td>.92 (.84 - .96)</td>
</tr>
<tr>
<td><strong>Varies type of grasp</strong></td>
<td>1.0</td>
<td>.94 (.88 - .97)</td>
</tr>
<tr>
<td><strong>Releases</strong></td>
<td>.79 (.53 - .90)</td>
<td>.96 (.91 - .98)</td>
</tr>
<tr>
<td><strong>Moves fingers</strong></td>
<td>.81 (.66 - .90)</td>
<td>.98 (.96 - .99)</td>
</tr>
<tr>
<td><strong>Grip force regulation</strong></td>
<td>.81 (.66 - .90)</td>
<td>.92 (.85 - .96)</td>
</tr>
<tr>
<td><strong>Manipulates</strong></td>
<td>.96 (.92 - .98)</td>
<td>.90 (.81 - .95)</td>
</tr>
<tr>
<td><strong>Coordinates</strong></td>
<td>.75 (.52 - .87)</td>
<td>.90 (.80 - .95)</td>
</tr>
<tr>
<td><strong>Orients objects</strong></td>
<td>.83 (.70 - .91)</td>
<td>.82 (.66 - .91)</td>
</tr>
<tr>
<td><strong>Proceeds</strong></td>
<td>.78 (.62 - .88)</td>
<td>.84 (.68 - .92)</td>
</tr>
<tr>
<td><strong>Flow in bimanual task performance</strong></td>
<td>.77 (.61 - .87)</td>
<td>.95 (.89 - .97)</td>
</tr>
<tr>
<td><strong>For total measure in AHA-units, ICC (95% CI)</strong></td>
<td>.97 (.94 - .98)</td>
<td>.99 (.98 - .99)</td>
</tr>
</tbody>
</table>

**Measurement error**

<p>| | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>SEM (AHA-units)</strong></td>
<td>2.3</td>
<td>1.6</td>
</tr>
<tr>
<td><strong>SDC (AHA-units)</strong></td>
<td>4.5</td>
<td></td>
</tr>
<tr>
<td><strong>Group mean difference (SD) (AHA-units)</strong></td>
<td>-.47 (3.3)</td>
<td>-.55 (2.3)</td>
</tr>
</tbody>
</table>

Ad-AHA, Assisting Hand Assessment in adolescents; ICC, intraclass correlation coefficients; SEM, standard error of the measurement; SDC, smallest detectable change.
Table II. Agreement between raters and alternate test forms, stability across time and smallest detectable change

<table>
<thead>
<tr>
<th>Alternate form</th>
<th>Alternate form</th>
<th>Alternate form</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ad-AHA Board Game vs. School-Kids AHA</td>
<td>Ad-AHA Board Game vs. Ad-AHA Present</td>
<td>Ad-AHA Board Game vs. Ad-AHA Sandwich</td>
</tr>
<tr>
<td>n=30</td>
<td>n=28</td>
<td>n=29</td>
</tr>
</tbody>
</table>

<p>| | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>.90 (.81 -.95)</td>
<td>.88 (.76 -.94)</td>
<td>.81 (.62 -.91)</td>
</tr>
<tr>
<td>.89 (.78 -.95)</td>
<td>.78 (.58 -.89)</td>
<td>.91 (.82 -.96)</td>
</tr>
<tr>
<td>.90 (.79 -.95)</td>
<td>.96 (.92 -.98)</td>
<td>.91 (.81 -.96)</td>
</tr>
<tr>
<td>.95 (.89 -.97)</td>
<td>1.0</td>
<td>.94 (.87 -.97)</td>
</tr>
<tr>
<td>.93 (.86 -.97)</td>
<td>.92 (.83 -.96)</td>
<td>.91 (.81 -.96)</td>
</tr>
<tr>
<td>.86 (.72 -.93)</td>
<td>.94 (.87 -.97)</td>
<td>.76 (.47 -.89)</td>
</tr>
<tr>
<td>.98 (.97 -.99)</td>
<td>.99 (.97 -.99)</td>
<td>.89 (.78 -.95)</td>
</tr>
<tr>
<td>.96 (.91 -.98)</td>
<td>.93 (.86 -.97)</td>
<td>1.0</td>
</tr>
<tr>
<td>.97 (.93 -.99)</td>
<td>.88 (.77 -.94)</td>
<td>.90 (.80 -.95)</td>
</tr>
<tr>
<td>1.0</td>
<td>.90 (.80 -.95)</td>
<td>.96 (.91 -.98)</td>
</tr>
<tr>
<td>.96 (.93 -.98)</td>
<td>.92 (.82 -.96)</td>
<td>.94 (.87 -.97)</td>
</tr>
<tr>
<td>.96 (.91 -.98)</td>
<td>.93 (.84 -.96)</td>
<td>.93 (.85 -.97)</td>
</tr>
<tr>
<td>.98 (.96 -.99)</td>
<td>.91 (.82 -.96)</td>
<td>.95 (.90 -.98)</td>
</tr>
<tr>
<td>.96 (.92 -.98)</td>
<td>.97 (.93 -.99)</td>
<td>.89 (.76 -.95)</td>
</tr>
<tr>
<td>1.0</td>
<td>.94 (.88 -.97)</td>
<td>.96 (.91 -.98)</td>
</tr>
<tr>
<td>.96 (.91 -.98)</td>
<td>1.0</td>
<td>.74 (.47 -.88)</td>
</tr>
<tr>
<td>.98 (.95 -.99)</td>
<td>1.0</td>
<td>.94 (.87 -.97)</td>
</tr>
<tr>
<td>1.0</td>
<td>.90 (.79 -.95)</td>
<td>.92 (.84 -.96)</td>
</tr>
<tr>
<td>.97 (.94 -.99)</td>
<td>.91 (.81 -.95)</td>
<td>.96 (.91 -.98)</td>
</tr>
<tr>
<td>.98 (.95 -.99)</td>
<td>1.0</td>
<td>.97 (.93 -.99)</td>
</tr>
<tr>
<td>.99 (.98 -.99)</td>
<td>.99 (.95-99)</td>
<td>.99 (.98 -.99)</td>
</tr>
</tbody>
</table>

<p>| | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1.4</td>
<td>1.8</td>
<td>1.5</td>
</tr>
<tr>
<td>3.8</td>
<td>5.0</td>
<td>4.1</td>
</tr>
<tr>
<td>-.23 (1.9)</td>
<td>1.5 (2.1)</td>
<td>.59 (2.1)</td>
</tr>
</tbody>
</table>
Chapter 3

Discussion

This study demonstrated excellent interrater and test–retest reliability of the AHA measures for adolescents with unilateral CP. The small test–retest SEM values indicate that measurement error due to random variance is acceptably low, taking into account not only the random error variance, but also possible rater and person inconsistency. Agreement between performance scores on the Ad-AHA Board Game and the School-Kids AHA was strong, which reflects excellent alternate-form reliability. Individual scores on the School-Kids AHA and the Ad-AHA Board Game were directly comparable, which makes it possible to use both test activities interchangeably for children aged 10 to 13 years. The alternate-form reliability was also good for adolescents (age 13–18y) performing three different activities. For interpretation of change knowledge about the size of the SDC is important. This study showed the SDC to be 4.5 AHA-units. Thus, a change of 5 AHA-units or more does (with 95% certainty) reflect a real change exceeding any random measurement error. This corresponds well with the SDC reported for the Kids-AHA version 4.4 (age 1y 6mo–12y).18

We calculated the SEM and SDC values of the alternate tests to determine whether the test activities can be used interchangeably. The higher the reliability, the lower the SEM, and the greater the confidence in the accuracy of an individual's test score. Because alternate-form reliability includes both the consistency of assisting hand use in different test activities as well as the consistency of performance across time, we expected that the estimated reliability coefficients would be somewhat lower and the SDC higher in comparison to test–retest values. We found a slightly higher SDC (5.0 AHA-units) for the Ad-AHA Board Game versus the Ad-AHA Present. The SDC for the Ad-AHA Board Game versus the School-Kids AHA was even lower (3.8 AHA-units), indicating that consistent performance is expected also when switching from the test activity appropriate for 10- to 12-year-old children (School-Kids Board Games) to the test activities appropriate for 13- to 18-year-old adolescents.

The different test activities of the Ad-AHA are important tools for observing and evaluating bimanual hand use in adolescents. The AHA 5.0 scoring system yielded consistent results even when the activities were performed in different ways due to cultural differences. For example, in the Sandwich task, some adolescents cut a slice of cheese with a knife (Australia) and some with a cheese slicer (the Netherlands and Sweden). There were also differences regarding the type of bread, spread, plastic bag closing devices, or wrapping paper and gift ribbons that were used in the different countries. However, these differences were discussed with participating therapists and it was ascertained that the tasks elicited bimanual actions sufficiently to make observation and scoring of AHA test items possible.

A limitation of this study might be the relatively small sample sizes to evaluate different types of reliability. It is generally accepted that at least 30 cases are needed for a reliable estimate of measurement error.19 Our samples for the alternate-form reliability in adolescents were somewhat smaller than the recommended sample size but it seems unlikely that this has affected the accuracy of the reliability indices. A limitation for the generalizability may be the fact that in our study, assessors had several years of experience with the AHA (one of them is an AHA teacher), which may have led to smaller measurement error. However, the test administration and AHA 5.0 scoring are highly standardized with precise and clear instructions specified in the AHA manual, which enhance test reliability. To evaluate test–retest reliability, the assessor was blind to previously obtained AHA scores.
and none of the participants received any form of intensive therapy between the two test occasions. This study has demonstrated that the SDC of 5 AHA-units can be considered as a value of a change beyond measurement error, which is useful for interpreting effects of interventions. The highly acceptable SDC of 5 AHA-units (on the 0–100 scale) and the fact that the AHA 18–18 can distinguish nine significantly different levels of ability indicate that the test has a high potential to be responsive to change. However, research is also needed to investigate the minimal clinically important difference. The minimal clinically important difference is the smallest change in scores that would be considered important or clinically meaningful to individuals. This may be challenging for several reasons, for example small changes (even less than the SDC of 5 AHA-units) may be very meaningful for one child or adolescent but not for another, and the size of a meaningful change may vary across different levels of ability. For example, for a person with severely limited function of the affected hand, a change from a score of 1 (‘does not do’) to a score of 2 (‘inefficient’) on a few easy items like ‘Amount of use’, ‘Holds’, ‘Stabilizes by weight or support’, may imply a big difference making it possible to start using the two hands together at all. This small change may open up a possibility to perform some bimanual tasks, even if it is with difficulty. On the other end of the ability scale, a person already scoring 3 (somewhat effective/almost normal performance) on the more difficult items like ‘Manipulates’, ‘Varies type of grasp’, ‘Readjusts grasp’, may not encounter scores of 4 (effective) as a clinically important difference since already a score of 3 indicates a high ability and good assisting hand function.

To report outcomes and evaluate change, interval level measures are recommended. Therefore the AHA raw-scores are converted into a logit-based 0 to 100 AHA-unit scale (AHA-units). The logarithmic transformation ‘stretches’ the scale in the lowest and the highest ends and renders a scale with equal size of measure intervals along the whole range. Thus, this scale makes change comparable for persons at different starting points along the scale. It also implies that, for example, a change of 1 AHA raw-score in persons with an extreme low ability (22 raw scores) results in a change of 4 AHA-units. On the other hand, a change of 1 AHA raw-score in persons with an average ability (about 40 raw scores) may result in a change of only 1 AHA-unit. For clinical practice and research, it is important to know that a small change in raw-scores in individuals with extreme low or high ability may imply a big (significant) change in measures of the quality of use of the affected hand during bimanual performance, and the interval level AHA-units is the fairest measure of change.

**Conclusions**

This study showed that the Ad-AHA has good-to-excellent interrater and test–retest reliability in adolescents with unilateral CP. Test scores for the alternate forms were comparable, showing that reliable AHA scores can be generated by using different age-appropriate tests situations for children and adolescents aged 18 months to 18 years. The excellent measurement agreement means that the different test activities can be used to evaluate change over time, with a score change of 5 AHA-units indicating a change beyond measurement error with 95% certainty, using the AHA 5.0 scale.
Acknowledgements

The authors thank the adolescents and their caregivers who provided data for the analyses and the therapists of the participating rehabilitation centres and departments in the Netherlands (Adelante Paediatric Rehabilitation Centre, Valkenburg; De Vogellanden Rehabilitation Centre, Zwolle; Sint Maartenskliniek Rehabilitation Centre, Nijmegen; De Trappenberg Rehabilitation Centre, Huizen; Sophia Rehabilitation Centre, Delft; Department of Rehabilitation Medicine, VU University Medical Centre, Amsterdam and Department of Rehabilitation, Academic Medical Centre, Amsterdam); Australia (Queensland Cerebral Palsy and Rehabilitation Research Centre); and Sweden (Neuropediatric Unit, Department of Women’s and Children’s Health, Karolinska Institutet, Stockholm). The authors gratefully acknowledge financial support from Stichting Rotterdam Kinderrevalidatie Fonds Adriaanstichting and Johanna KinderFonds (the Netherlands) and Stiftelsen Frimurare-Barnhuset, the Strategic Research Funding for Health Care Science (SFO-V), and the Sunnerdahl foundation (Sweden).

LK-S is involved in the company Handfast AB (Inc.), which supplies training courses and test material for the Assisting Hand Assessment. KB is an AHA teacher and organizes AHA courses in the Netherlands. Neither company has been involved in or funded any part of the research. The remaining authors have stated that they had no interests which might be perceived as posing a conflict or bias.

References


