## Global Health

# Validation of the pediatric Radboud Dysarthria Assessment

Marieke Ruessink[a], Lenie van den Engel-Hoek[b], Marjo van Gerven[b,*], Bea Spek[c],
Bert de Swart[b] and Johanna Kalf[b]

[a]*Sint Maartenskliniek, Department of Rehabilitation, Nijmegen, The Netherlands*
[b]*Radboud University Medical Center, Donders Institute for Brain, Cognition and Behaviour,
Department of Rehabilitation, Nijmegen, The Netherlands*
[c]*Department of Clinical Epidemiology, Biostatistics and Bioinformatics, Amsterdam University Medical Centers,
University of Amsterdam, Amsterdam, The Netherlands*

**Abstract**.

**PURPOSE:** The Radboud Dysarthria Assessment (RDA) was published in 2014. Adaptation into a pediatric version (p-RDA) was required because of relevant differences between children and adults. The purpose of this study was to assess the feasibility of the p-RDA and to test intra-rater and inter-rater reliability as well as the validity of the two severity scales (function and activity level).

**METHODS:** Video recordings were made of 35 participants with (suspected) dysarthria (age 4 to 17 years) while being assessed using the p-RDA. Intra-rater reliability was assessed by one, and inter-rater reliability by two experiments using the Intraclass Correlation Coefficient (ICC). Validity of the severity scales was tested by correlating the consensus scores with the independently rated scores on four communication scales, three mobility scales, and one self-care scale using Spearman correlation coefficients ($r_s$).

**RESULTS:** The assessment was applicable for 89% of the tested sample, with good intra-rater and inter-rater reliability (ICC = 0.88–0.98 and 0.83–0.93). The p-RDA severity scales (function and activity level) correlated from substantially to strongly with the communication scales ($r_s$ = 0.69–0.82 and 0.77–0.92) and self-care scale ($r_s$ = 0.76–0.71) and correlated substantially with the mobility scales ($r_s$ = 0.49–0.60).

**CONCLUSION:** The feasibility, reliability and validity of the p-RDA are sufficient for clinical use.

Keywords: Childhood dysarthria, dysarthria assessment, severity, reliability, validity

## 1. Introduction

Dysarthria is a motor-speech disorder that may affect articulation, resonance, phonation, prosody, and respiration. It can occur in both children and adults and is caused by a central or peripheral neurological disorder [1]. Pediatric dysarthria may have profound effects on children's ability to communicate, which is essential for a child's quality of life, development, social interaction, and participation [2, 3]. The prevalence of childhood dysarthria is understudied; however, data suggest that 33–63% of children with cerebral palsy (CP) experience a motor speech impairment [4–6].

Proper assessment is essential in order to identify and understand dysarthria, yet in the Netherlands a standardized and valid assessment for diagnosing

*Corresponding author: Marjo van Gerven, Radboud university medical center, Donders Institute for Brain, Cognition and Behaviour, Department of Rehabilitation, P.O. Box 9101, Internal Code 898, 6500 HB Nijmegen, The Netherlands. Tel.: +31 24 3668185; E-mail: marjo.vangerven@radboudumc.nl.

childhood dysarthria was insufficient and there were no valid assessments available in other languages. Furthermore, there was a strong need among speech-language therapists (SLTs) for standardized and valid methods for diagnosing dysarthria based on consensus [7]. For adult patients, this need has been met by the (Dutch) Radboud Dysarthria Assessment (RDA) published in 2014 [8, 9]. This is an observational assessment instrument based on two speech tasks and four speech-related tasks. The two speech tasks comprise spontaneous speech during a short conversation (about any topic) and reading a standardized text aloud. The four others are speech-related maximum performance tasks: maximum repetition rate (diadochokinesis), maximum frequency range (from the lowest to the highest pitch), maximum phonation volume, and maximum phonation duration (on vowel [a]). Spontaneous speech and reading are considered to represent daily communicative functioning. The maximum performance tasks were included for clinicians to challenge and observe the various aspects of speech in detail [9]. Based on the interpretation of these speech performances, clinicians are advised to use the Mayo Clinic classification [10] when deciding on the dysarthria type and judging the severity of the dysarthria on two six-point ordinal scales, based on Therapy Outcomes Measures by Enderby & John [11]. One scale assesses the severity of the dysarthria at function level, the other one is a scale for activity level (effectiveness of verbal communication), see Fig. 1. These severity scales were described for adult dysarthria by Knuijt et al. [9]. The severity score is an overall score that reflects the most severely affected aspect of speech production (articulation, phonation, or resonance). Clinicians are asked to allocate an overall severity score (function and activity level) of the dysarthria based on the various tasks. Both scales proved to be reliable and valid in the RDA [9].

Although this assessment consists of common speech tasks, adaptation into a pediatric version was necessary due to relevant differences between children and adults. Firstly, adult dysarthria generally involves a speech system that once was intact, whereas a childhood dysarthria usually involves a developing motor, cognitive, and linguistic system. Secondly, children with CP or other neurological diseases often also have phonological and language deficits in addition to their neurologic speech (and sometimes cognitive) deficits. Consequently, atypical development in one domain may affect the others [12]. Furthermore, the neural basis, the cause of dysarthria, and the speech characteristics in children

and adults often differ, meaning an adult-based neurobehavioral classification system cannot be validly applied to children [1, 13]. Together, these differences complicate the interpretation of pediatric dysarthria. Finally, the assessment needed to be feasible for children in the age range of 5 to 18 years. For these reasons, the authors constructed a pediatric version: the pediatric Radboud Dysarthria Assessment (p-RDA).

The basic structure of the adult version was maintained for the p-RDA, and consisted of two observational tasks (word and sentence repetition or a reading task, and a conversation about a familiar topic e.g., favorite play), four maximum performance tasks, and two severity scales (Fig. 1). However, in the manual the task instructions were rewritten for use with children and two easier reading tasks with different reading levels were added as well. The selection of reading passages was based on a Dutch system for the determination of levels of difficulty (Dutch AVI-level M4 and M6). The AVI-level M4 text consists of 101 words; the M6 text of 194 words. For the oldest children and young adults who were able to read proficiently, the reading task of the RDA (adult version) which consists of 520 words was added [9]. In addition, for children who were unable to read, a list of words and sentences was added for them to imitate. The process of word selection was based on the picture-naming task of the Computer Articulation Instrument (CAI) [14]. A total of 60 words were included in the CAI, which incorporates the full body of vowel, consonant, cluster, and syllable structure combinations of the Dutch language. The syllable shapes of the target words varied from simple to more complex. With respect to the p-RDA, 20 one-syllable words and 10 two or three-syllable words were selected for the youngest children. For the older children and young adults, 20 two and three-syllable words were added to the list. In addition, children were asked to repeat 10 sentences, 2 without nasal consonants, 2 with nasal consonants, and 6 combined sentences.

This assessment was intended to be performed in about 15 to 20 minutes for children older than 5 years. The final version was discussed, adapted, and approved in a consensus meeting with Dutch SLTs experienced in childhood dysarthria.

The aim of this study is (1) to assess the feasibility of the p-RDA in children with neurological disorders (2) to test the intra-rater and inter-rater reliability and (3) to assess the construct validity (demonstration that the test is actually measuring the intended

**Severity scale at function level:**

0.  No dysarthria.

1.  Minimal dysarthria: minimal problems with articulation, resonance, phonation or respiration.

2.  Mild dysarthria: tonus, range of motion or speed of articulation movements are mildly affected, leading to mildly affected consonants and/or vowels. Mild problems with resonance, phonation or respiration.

3.  Mild/severe dysarthria: tonus, range of motion or speed of articulation movements are clearly affected, leading to deviant consonants and/or vowels. Clearly affected resonance, phonation or respiration.

4.  Severe dysarthria: tonus, range of motion or speed of articulation movements are severely affected, leading to severely affected consonants and/or vowels. Severely affected resonance, phonation or respiration. Very slow speech with only a few syllables per breath group.

5.  Very severe dysarthria/anarthria: articulation movements are almost impossible, leading to mainly open vowels and very severely distorted consonants. Very severely affected resonance, phonation (aphonic) or respiration (almost no breath support).

**Severity scale at activity level:**

0.  Effective communication: taking into account both chronological and developmental age as well as cultural aspects, communication is effective in all situations.

1.  Effective despite small problems: can be understood most of the time by any listener despite communicative irregularities.

2.  Incidental repetitions: communication with familiar persons is effective. With some help, communication with unfamiliar persons may be effective. Capable of communicating beyond here and now. Occasional repetition required.

3.  Frequent repetitions: able to communicate basic needs. Frequent repetition required. Communication improves when using trained listeners and family members or in familiar settings.

4.  Communication with a known person: occasionally able to communicate basic needs with familiar persons or trained listeners in familiar contexts.

5.  No oral communication possible. No effective verbal communication possible.

Fig. 1. p-RDA severity scale regarding function level and activity level [9].

Table 1
Patient characteristics of the group children who were assessed with the p-RDA ($n = 35$)

| Characteristics | |
| --- | --- |
| Mean age (SD) (years;months) | 9;03 (3;02), range 4;03–17;10 |
| Sex (boys–girls) | 20 (57%)–15 (43%) |
| | *N* |
| Main diagnosis | |
| *Central neurological disorder (CND)* | |
|   Cerebral Palsy | 20 |
|   Epilepsy - Hemispherectomy | 1–2 |
|   Syndrome | 2 |
|   Medulla blastoma | 1 |
| *Peripheral neurological disorder (PND)* | |
|   Myotonic dystrophy | 2 |
|   Dystrophy type Duchenne | 2 |
|   Mitochondrial Encephalomyopathy | 1 |
| *Neurodegenerative disease (NDD)* | |
|   Ataxia Telangiectasia | 1 |
| *Unknown* | 3 |
| *Other oral motor, speech or language difficulties\** | |
|   Yes–No–Unknown | 21–5–9 |
| *Native speaker Dutch* | |
|   Yes–No | 26–3 |
|   Unknown | 6 |

*Other oral motor, speech and/or language difficulties as noted by the children's SLT in their file using standardized tests.

construct) by correlating the severity scales of function and activity level against a set of validated scales that measure the same or a proximate construct. For the construct validity it was hypothesized that the correlations of the p-RDA severity scales with the scales in the communication domain would be stronger than the correlations of the p-RDA severity scales with the scales in the domains of mobility and self-care.

## 2. Method

### 2.1. Participants

For the validation of the severity scale, 35 children ages 5 to 18 years with (suspected) acquired or congenital dysarthria were recruited, see Table 1. The lower limit of 5 years was selected because speech-language development is mostly complete at this age.

The suspected dysarthria diagnosis was based on information provided by the SLT of the child on their prior speech-language therapy and established neurological diagnosis. There were no exclusion criteria except for the ability to understand the

instructions and to vocalize or speak. All participants received outpatient-rehabilitation treatment at Klimmendaal Rehabilitation Specialists in Arnhem or the Radboud University Medical Centre, Department of Rehabilitation in Nijmegen (the Netherlands). Assessments were conducted between January and June 2015 after approval by the Medical Ethics Committee (NL51490.091.14). Children aged 11 years or younger gave verbal consent and their parents or caregivers provided written consent for participation in the study. For those older than 12, written consent was obtained from the children and their parents or caregivers.

### 2.2. Measurement procedures

All children were assessed using the p-RDA and this was recorded on video. The assessment was conducted during daytime under standardized conditions according to preparation (optimal body position of the child, a room without ambient noise), camera position (frontal at eye level), and position of an external microphone (30 cm distance from the child's mouth). The assessments were recorded on digital video and stored on a secured hard disk. Other tasks were recorded with a Tascam DR-05 and the maximum performance tasks were measured with a Db-meter (Voltcraft SL-100).

Thirty-two children were assessed by the first author (AUTHOR 1) and three children by AUTHOR 1 and their own SLT together.

### 2.3. Feasibility

During data collection, three aspects were observed and registered: whether the children could perform all tasks, whether they were able to perform the tasks in one attempt, and the duration of the assessments.

### 2.4. Reliability

In order to assess inter-rater reliability of the severity scales, two experiments were established. First, two highly experienced SLTs (AUTHORS 2 AND 3) independently rated the p-RDA severity scales, based on 35 videos (inter-rater reliability). The raters were asked to give an overall severity score of the dysarthria by using the p-RDA-severity scale at function level and activity level based on the speech tasks and the maximum performance tasks.

Table 2
Scales selected to determine construct validity of the p-RDA severity scale

**Selected scales**

| Domain | For | Correlating scale | To be reported by |
|---|---|---|---|
| Communication | All children | Intelligibility in Context Scale: Dutch (ICS)[15] 5 = high intelligibility, 1 = low intelligibility | Parents or caregivers |
| | All children | Communication Function Classification System (CFCS)[16] 1 = effective communication both verbal and non-verbal, 5 = seldom effective communication both verbal and non-verbal | SLT |
| | All children | Viking Speech Scale (VSS)[17] 1 = speech not affected, 4 = no understandable speech | Health professional |
| | All children | Capacity Profile Voice and Speech (CAP)[18] 0 = understandable speech should be possible, 5 = no communication possible | Rehabilitation specialist |
| Mobility | Children with central nervous system disease | Manual Ability Classification System (MACS)[19] 1 = handles objects easily, 5 = does not handle objects | Occupational therapist |
| | | Gross Motor Function Classification System (GMFCS)[21] 1 = climb stairs without limitations, 5 = wheelchair bounded | Pediatric physical therapist |
| | Children with peripheral nervous system disease | Vignos and Brooke classification (V&B)[20] Upper extremities: 1 = no limitations, 6 = no functional use of hands Lower extremities: 1 = climbing stairs without assistance, 10 = bed ridden | Pediatric physical therapist |
| Self-care | All children | Functional Oral Intake Scale (FOIS)[22] 7 = oral diet with no restriction, 1 = nothing by mouth | Parents or other caregivers |

The activity scale was based on rater judgements of the intelligibility of word and sentence repetition or a reading task and spontaneous speech during a conversation. In addition, clinicians working with a pediatric dysarthric population may also come across dysarthria in combination with other speech-sound disorders (SSD), such as phonological problems and/or childhood apraxia of speech (CAS). However, raters took this into account in their judgements.

After the independent scores had been allocated by each author, a consensus score was needed for validation (in next paragraph). This score was reached by having the two raters view the videos together and agree on the appropriate scores.

For the second experiment 17 SLTs, all working in pediatric rehabilitation, received a two-hour training to become familiar with the p-RDA. The training consisted of the following parts: (1) theoretical information of the p-RDA; (2) introduction of the severity scales and team-wise severity scoring of 2 patients; (3) evaluation of this scoring. After the training, the 17 SLTs independently rated the p-RDA severity scale based on two other selected videos (inter-rater reliability).

In order to assess intra-rater reliability, the 17 trained SLTs were asked to rate the videos at two different time points separated within 8–11 weeks. Only 8 SLTs were able to score the videos twice (missing 9 scores).

### 2.5. Validity

In order to estimate the validity of the severity scales, the consensus scores were compared with other scales in use for children with central or peripheral disorders, either speech related (convergent validity) or non-speech related (divergent validity).

The process of selecting convergent and divergent scales was as follows. A total of 104 scales and tests were screened based on five criteria: patient-important outcome, practical applicability (burden of use for the child or the parents, profession of the assessor), suitable age range, established validity, and availability in Dutch. This screening identified 68 scales, of which 12 best met the five criteria. To further narrow down the number of scales, patient-important outcome and practical applicability were prioritized, which resulted in 4 scales that were speech related: Intelligibility in Context Scale: Dutch (ICS [15]), Communication Function Classification System (CFCS [16]), the Viking Speech Scale (VSS [17]), and the Capacity Profile Voice and Speech (CAP [18]), which are shown in Table 2 (communication domain). These scales were expected to clearly correlate with the p-RDA severity scales and thus all four were deemed applicable to dysarthric children.

Four non-speech scales were selected, which were expected to have a much weaker correlation with the p-RDA scales. For children with a central nervous

system disease (like cerebral palsy) the common classification systems Manual Ability Classification System (MACS [19]) and Gross Motor Function Classification System (GMFCS [20]) were used. These mobility scales have been shown to moderately correlate with the CFCS [21]. Consequently, a similar correlation with the p-RDA scales was anticipated. For children with peripheral nervous system disease, the Vignos and Brooke classification (V&B scale [22]) was used. In conclusion, for all children, it was anticipated that the utility of the Functional Oral Intake Scale (FOIS [23]) as another non-speech scale (self-care domain) would be limited.

Within three weeks of recording the assessment with the p-RDA on video, the selected convergent and divergent scales were rated by health professionals or parents and caregivers involved in the care of the children (in Table 1). The health professionals or parents and caregivers were blinded for the other scales and the consensus scores of the p-RDA.

### 2.6. Data analysis

For feasibility, descriptive statistics were used. Inter-rater reliability was evaluated by using the intraclass correlation coefficient (ICC) two-way random for absolute agreement, single measure (ICC2,1). The intra-rater reliability was assessed using an ICC two-way mixed for consistency and single measure (ICC3,1) accepting 0.70 as a minimum for both coefficients [24].

To estimate construct validity, p-RDA severity scales were correlated with the selected scales by calculating Spearman correlation coefficients ($r_s$) because all scales were ordinal. Correlations $< 0.30$ were considered as weak, between 0.30 and 0.70 as substantial, and $\geq 0.70$ as strong [25] and 95% confidence intervals were calculated for all reliability and correlation coefficients.

The authors hypothesized substantial to strong correlations of the p-RDA severity scales with all speech scales (communication domain), because these are all intended to quantify the same construct (severity of disordered speech). For this dysarthric population, it was also hypothesized that there would be substantial correlations of the p-RDA severity scales with the non-speech scales (mobility and self-care domains), but much lower than compared to the speech scales.

Missing items were handled with complete case-only analysis. SPSS 21.0 for Windows (SPSS Inc., Chicago, IL, USA) was used for all calculations, and two-tailed $p$-values $\leq 0.05$ were considered to be statistically significant. Point estimates of the ICCs are presented with their 95% confidence intervals.

## 3. Results

Thirty-five participants agreed to participate in this study. Median age was 9 years and 3 months (interquartile range 4 years and 3 months–17 years and 10 months). Patient characteristics are summarized in Table 1, showing that in this cohort cerebral palsy was by far the most frequent disease. SLTs of 21 children noted comorbid oral motor, speech and/or language difficulties in their file based on regular assessments (5 children had no comorbid difficulties, 9 children did not have comorbid difficulties noted). Differential diagnoses were part of the clinical reasoning process of the SLT and were made based on diagnostic criteria described in studies such as Forrest (2003) [26], Shriberg and Kwiatkowski (1994) [27], and Morgan and Liégeois (2010 and 2012) [1, 13]. These SLTs also applied their usual assessment tools such as the Metaphon screening, the Dutch assessment of Hodson and Paden, or the assessment of the Nuffield Dyspraxia program. The difficulties were: drooling ($n = 1$), dysphagia ($n = 4$), developmental language delay and phonetic speech disorder ($n = 3$), developmental language delay and dysphagia ($n = 2$), childhood apraxia of speech ($n = 3$), developmental language delay and hearing loss ($n = 1$), childhood apraxia of speech and phonological problems ($n = 1$), childhood apraxia of speech and developmental language delay ($n = 1$), phonological problems ($n = 3$), and childhood apraxia of speech, developmental language delay, and phonological problems ($n = 2$).

The distribution of the scores on all scales are presented in Table 3. A few missing values had to be recorded: one in the p-RDA severity scale at function level and two in the activity scale because of profound comorbidity (phonological disorder and/or childhood apraxia of speech (CAS)), as well as a few in the ICS, CFCS, VSS, CAP, MACS, GMFCS and FOIS due to non-response (see Table 3 for details).

### 3.1. Feasibility

With all participants, the SLTs were able to do the assessment within 20 minutes; all children except 4 (due to very severe dysarthria with characteristics of involuntary movements) were able to complete the maximum performance tasks. The assessment was applicable for 89% of the tested sample ($n = 31$).

Table 3
Scores of the scales (data of 35 participants)

| | Median | Range |
|---|---|---|
| **Severity p-RDA scales** | | |
| p-RDA severity scale at function level (0 = no dysarthria, 5 = very severe)[1] | 2 | 0–5 |
| p-RDA severity scale at activity level (0 = effective communication, 5 = no oral communication possible)[2] | 3 | 0–5 |
| **Convergent and divergent scales** | | |
| *Communication* | | |
| Intelligibility in Context Scale: Dutch (1 = low intelligibility, 5 = high intelligibility)[3] | 3.5 | 1–5 |
| Communication Function Classification System (1 = effective communication both verbal and non-verbal, 5 = rarely any effective communication neither verbal nor non-verbal)[4] | 3 | 1–4 |
| Viking Speech scale (1 = speech not affected, 4 = no understandable speech)[5] | 3 | 1–4 |
| Capacity Profile Voice and speech (0 = understandable speech should be possible, 5 = no communication possible)[6] | 1 | 0–4 |
| *Mobility* | | |
| Manual Ability Classification System (1 = handles objects easily, 5 = does not handle objects)[7] | 3 | 1–5 |
| Gross Motor Function Classification System (1 = climb stairs without limitations, 5 = wheelchair bound)[8] | 2 | 1–5 |
| Vignos and Brooke Classification upper extremities (1 = no limitations, 6 = no functional use of hands) | 1 | 1–2 |
| Vignos and Brooke Classification lower extremities (1 = climbing stairs without assistance, 10 = bedridden) | 2 | 1–9 |
| *Self-care* | | |
| Functional Oral Intake Scale (7 = oral diet with no restriction, 1 = nothing by mouth)[9] | 7 | 3–7 |

The higher the score, the higher the level of restrictions, except for the ICS and FOIS scale. [1]One missing value (3%), because SLTs could not reliably judge the dysarthria severity due to profound hearing loss with cochlear implants at both sides; this child was unable to understand all the commands for the tasks, especially the maximum performance tasks. [2]Two missing values (6%), because SLTs could not reliably judge the dysarthria severity due to profound comorbidity (severe phonological problems and childhood apraxia of speech); these children proved unable to react during the conversation, which is needed to judge oral communication. [3–7]Missing values because of non-response from health workers: 6 (17%) at the ICS, 5 (14%) at the CFCS, 5 (14%) at the VSS, 6 (17%) at the CAP, 5 (14%) at the GMFCS, 6 (17%) at the MACS, 6 (17%) at the FOIS.

In this study, 26 participants completed the repetition task and 9 participants completed the reading task.

### 3.2. Reliability and validity

In the first experiment with two raters, the ICCs for the severity scales at function and activity level were 0.83 (95% CI 0.62–0.92) and 0.86 (95% CI 0.75–0.93). The Bland-Altman plot (Fig. 2a) for the function scale reflects the differences between the two raters against the mean difference. The scores of the second rater were generally lower than those of the first rater, indicating a systematic error of 0.46 points (SD 0.82). All scores except one were between the lower and upper limit of ± 1.96. The Bland-Altman plot for the activity scale shows a systematic error of 0.03 points (SD 0.82) (Fig. 2b). Three points were outside the lower and upper limit of ± 1.96.

For the raters in the second experiment, the inter-rater reliability for the function scale was 0.91 (95% CI 0.64–1.0) and 0.93 for the activity scale (95% CI 0.69–1.0). The ICCs of the intra-rater reliability were 0.88 (95% CI 0.69–0.96) and 0.98 (95% CI 0.95–0.99) (see Table 4).

The correlation coefficients for the construct validity are adopted in Table 5. The function scale correlated substantially to strongly with all speech scales

(communication domain) with correlations coefficients ranging from $r_s = 0.69$ to –0.85, ICS $r_s = -0.85$ (95% CI –0.94 – –0,67), CFCS: $r_s = 0.69$ (95% CI 0.41–0.85), VSS: $r_s = 0.82$ (95% CI 0.62–0.92), and CAP: $r_s = 0.78$ (95% CI 0.54–0.90). The action scale showed similar correlations with the ICS and CAP, but higher correlations with the CFCS (0.87; 95% CI 0.71–0.94) and VSS (0.92; 95% CI 0.79–0.96).

The function scale also correlated substantially, albeit much lower than with the two mobility scales and with very wide confidence intervals: MACS: $r_s = 0.49$ (95% CI 0.07–0.76), GMFCS: $r_s = 0.49$ (95% CI 0.08–0.76), while their correlation with the activity scale was $r_s = 0.60$ and with confidence intervals of 0.21–0.82 and 0.22–0.82. No linear association was found with the V&B scale. In conclusion, both severity scales correlated strongly with the FOIS: $r_s = -0.76$ (95% CI –0.50 – –0.89), and $r_s = -0.71$ (95% CI –0.42 – –0.86).

## 4. Discussion

The p-RDA is a clinical tool uniquely designed to assess childhood dysarthria. As an adaptation of the original assessment for adults, the feasibility of the full assessment and the reliability of its severity scales

2a. Bland Altman plot for the dysarthria severity scale (function level)
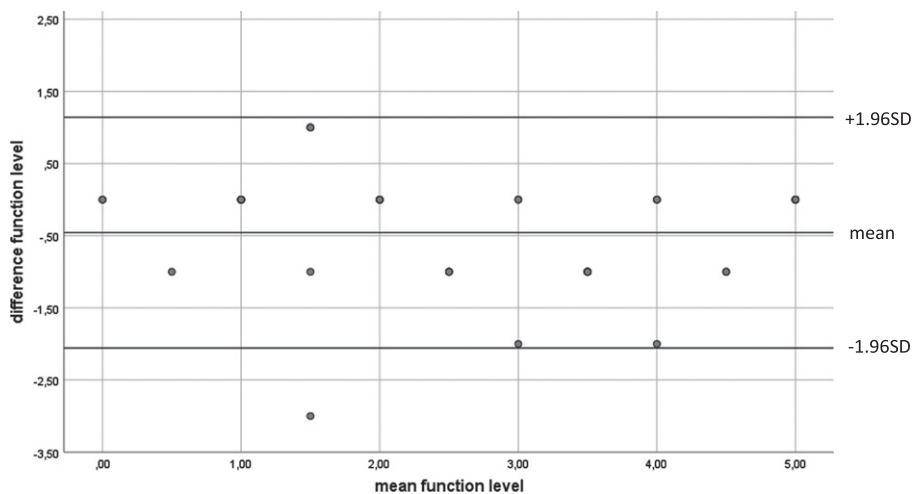


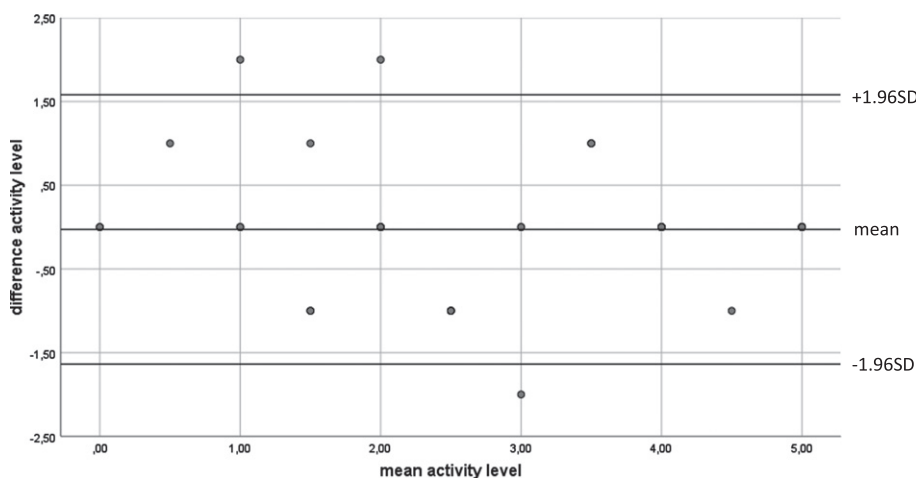2b. Bland Altman plot for the dysarthria severity scale (activity level)



Fig. 2. Bland Altman plots for the dysarthria severity scale.

Table 4
ICC-values with 95% confidence intervals of intra- and inter-rater reliability measures.
All participants participated in the 2-hour training.

| | Severity scale: function | Severity scale: activity |
| --- | --- | --- |
| Intra-rater | | |
| –8 raters, 2 videos[1] | 0.88 (95% CI 0.69–0.96) | 0.98 (95% CI 0.95–0.99) |
| Inter-rater | | |
| –2 raters, 35 videos (experiment 1) | 0.83 (95% CI 0.62–0.92) | 0.86 (95% CI 0.75–0.93) |
| –17 raters, 2 videos (experiment 2) | 0.91 (95% CI 0.64–1.0) | 0.93 (95% CI 0.69–1.0) |

[1]From the 17 raters, only 8 were able to score the videos a second time (missing 9).

are of the same standard as the clinimetric proper-
ties of the adult version, where ICCs of 0.81–0.85
were found [9]. However, it must be mentioned that
except for the intra-rater reliability of the activity
scale, the confidence intervals in the current study are
broad which indicates some uncertainty. In addition,
the construct validity of both the adult and pediatric
version was found to be good.

A clinical tool needs to be feasible regarding the
ability of participants to reasonably perform the tasks

Table 5
Correlation matrix of construct validity correlations

| Reported by | | Communication domain | | | | Mobility domain** | | Self-care domain |
|---|---|---|---|---|---|---|---|---|
| | | Parents | SLT | Health worker | Rehabilitation specialist | Occupational therapist | Pediatric physical therapist | Parents and SLT |
| | Scale | ICS[1] | CFCS[2] | VSS[3] | CAP[4] | MACS[5] | GMFCS[6] | FOIS[7] |
| p-RDA: function level* | | −0.85 | 0.69 | 0.82 | 0.78 | 0.49 | 0.49 | −0.76 |
| | | ($n = 29$) | ($n = 30$) | ($n = 30$) | ($n = 29$) | ($n = 23$) | ($n = 24$) | ($n = 29$) |
| 95 % CI | | −0.94 − −0.67 | 0.41–0.85 | 0.62–0.92 | 0.54–0.90 | 0.07–0.76 | 0.08–0.76 | −0.50 − −0.89 |
| p-RDA: activity level* | | −0.82 | 0.87 | 0.92 | 0.77 | 0.60 | 0.60 | −0.71 |
| | | ($n = 29$) | ($n = 30$) | ($n = 30$) | ($n = 29$) | ($n = 23$) | ($n = 24$) | ($n = 29$) |
| 95% CI | | −0.92 − −0.62 | 0.71–0.94 | 0.79–0.96 | 0.52–0.90 | 0.21–0.82 | 0.22–0.82 | −0.42 − −0.86 |

Communication domain · Reported on

[1]ICS = Intelligibility in Context Scale; [2]CFCS = Communication Function Classification System; [3]VSS = Viking Speech scale; [4]CAP = Capacity Profile Voice and speech; [5]MACS = Manual Ability Classification System; [6]GMFCS = Gross Motor Function Classification System; [7]FOIS = Functional Oral Intake Scale. *Data are presented as Spearman correlation coefficients. **No linear association was found between the p-RDA severity scales and the Vignos and Brooke scale ($n = 6$ children), and therefore correlations with the Spearman correlation coefficient could not be reported.

and the time it takes to complete them. Whereas occasionally it may be difficult to persuade children to do (unusual) speech tasks, all the children in this study cooperated without hesitation and the whole assessment could be performed in less than 20 minutes, which fits within the usual appointment time for one consultation. Although only four children with very severe dysarthria/anarthria (score 5) were not able to complete all tasks, their responses were relevant to observe and judge the severity of their dysarthria. In one function level and two activity level cases, judgement of the severity was difficult due to profound comorbidity, meaning that dysarthria was seen in combination with other speech problems. Although the authors recognize that some children in the research group had a combination of dysarthria and other speech problems (such as phonological problems and/or CAS), these children were included as it provided an opportunity to include a group of children that is a reflection of the complicated dysarthric pediatric population. In scoring, there was a difference in rating criteria between the severity scale at function level and the severity scale at activity level. Only the severity of the dysarthria was taken into account in the severity scale at function level. In this scale at activity level, all speech problems could lead to the situation of not being understood. Consequently, at this level it was not only the dysarthria that contributed to the level of severity.

All dysarthria severities were present in this consort, which is important for the clinical implementation because SLTs working in the field of pediatric rehabilitation face the full range of childhood dysarthrias.

A review by McCauley and Strand (2008) on standardized tests of nonverbal and speech motor performance in children included six tests [28]. They looked at reference values, criterion-referred measures, along with reliability and validity. Only the Verbal Motor Production Assessment for Children provided norms that were adequately described. The p-RDA provides added value by including the assessment of functional speech in combination with maximum performance tasks. This combination is vital in the assessment of dysarthria, because the five aspects of speech production (respiration, phonation, articulation, resonance, and prosody) interact with each other and can all be involved in dysarthria [8]. Maximum performance tasks of speech production examine the upper limits of speech motor performance and are considered useful to investigate the articulatory and phonatory-respiratory systems more independently than in spontaneous speech [29]. In addition, the psychometric evaluation of the p-RDA is extensively compared to current speech motor performance tests in children who often do not meet the criteria of relevant psychometric principles [28].

When describing a complex speech disorder such as pediatric dysarthria, a shared vision by SLTs on the severity levels is needed. The same descriptive scales (Fig. 1) were used as those in the adult version of the RDA, which means it is relevant, but at the same time not surprising that these analyses also showed intra-rater reliability scores well above 0.85: severity scale function level 0.88 (95% CI 0.69–0.96), severity scale activity level 0.98 (95%CI 0.95–0.99). Inter-rater reliability scores were also high: severity scale function level 0.83 (95% CI 0.62–0.92),

severity scale activity level 0.86 (95%CI 0.75–0.93). The ICCs of the agreements between the 17 raters was even above 0.90: severity scale function level 0.91 (95% CI 0.64–1.00), severity scale activity level 0.93 (95% CI 0.69–1.00) which could be explained by the fact that their judgement was limited to two videos (Table 4). The ICCs of the activity scales are slightly higher with narrower confidence intervals suggesting that it is easier to score the effectiveness of verbal communication than levels of speech performances. In addition, when a score depends on subjective judgement, familiarity with the rating scale is imperative for its reliability. That is why in this study the SLTs were trained by a two-hour video example. In order to assure a well-established scoring procedure, a manual (in Dutch) was published accompanying the instrument that explains in detail how to perform the assessment, scoring, and interpretation.

In order to organize a systematic validity testing of the severity scales, those in different domains were deliberately selected. Both the parents and the health professionals of the children could be asked to judge their motor functioning based on these scales, independent from each other. As expected, the correlations of the p-RDA severity scales with the those in the communication domain were stronger than the correlations of the p-RDA severity scales with the scales in the mobility domain. The scales that measure communication and speech (CFCS and VSS) showed higher correlations with the severity scale at activity level compared to the scale at function level (Table 5). The rating of a speech scale by persons other than trained SLTs would likely be closer to the activity scale for communicative effectiveness. This severity scale at activity level could also be scored by non-SLTs, whereas the severity scale on speech function requires detailed expertise in speech pathology. This could also explain the slightly lower reliability scores of the function scale.

The scales that measure mobility (MACS and GMFCS, but not V&B scale because of absence of linear association) showed moderate correlations and are again higher for the activity scale that rates communicative effectiveness, as the MACS and GMFCS are also scales for activity level. The moderate correlations between severity scales of the p-RDA and gross and fine motor skills are a reflection of the overall severity of the child's impairment and are in line with the results of the study of Hidecker et al. [21]. In addition, although motor speech disorders are related to other motor problems in these children, it is not

always in the same manner or of the same severity level [30].

Additionally, the domain of self-care (the FOIS) showed a strong correlation with the p-RDA severity scales. The FOIS intends to describe the type and amount of oral intake of food and liquid that a patient consumes on a daily basis [23]. The strong correlation between the FOIS and the severity scales of the p-RDA suggests that oral intake and childhood dysarthria are related, but does not imply a causal relationship. Studies suggest that there is no causal relationship because of the task-dependence assumption of laryngeal and orofacial muscles [31]. Overall, several correlations clearly show wide confidence intervals, indicating clinical uncertainty. More data should be collected before firm conclusions can be drawn.

A limitation of this study is the asymmetric distribution of etiologies. Rare dysarthrias may be more difficult to judge than common dysarthrias, but the assessment and its severity scales should be applicable to all dysarthrias. More than half (57%) of the cohort was represented by children with CP; nine other etiologies were present, suggesting that the high reliability and validity scores apply to all dysarthrias. However, future studies should be performed to confirm this reliability and validity in non-CP groups, and in younger children as well.

The sample size of the included children could be interpreted as another limitation. Therefore, use of the p-RDA in a larger group of children will help to confirm the findings of this research.

Another point of discussion is the use of speech characteristics when observing both spontaneous speech and interpreting the maximum speech performances. This serves to obtain a comprehensive view of the speech possibilities of children. In the adult assessment, these characteristics helped the SLT to diagnose the dysarthria using the Mayo Clinic classification [10]. However, in pediatric dysarthria, diagnosis and classification is not possible with this method because of the different causes of neurological damage in the setting of developing motor systems in children who may also have phonological and language deficits [32]. Comorbidity and comorbid speech disorders (for example childhood apraxia of speech) might affect the child's speech performance. However, in clinical care of pediatric dysarthria SLTs often come across various comorbidities. This makes the p-RDA even more clinically useful. In this study, scoring of the severity scales was constantly based on the overall performance. A more detailed assessment

(for example dysarthric and apraxic components) could help the SLT to provide tailored therapies. The classification of dysarthria subtypes in children requires reconsideration [13]. One intermediate step could be the determination of the most influential characteristic for intelligibility in pediatric dysarthria. The scoring system used and described in the manual of the p-RDA could be used for this purpose, and allows users to describe the main characteristics (for example hyperkinetic, hypokinetic, weak) of breathing, phonation, articulation, nasal resonance, and prosody. Future studies are needed to test such a format in order to develop a clinical approach on how to best use test performances to identify the most relevant speech characteristics of a child.

Where it was previously uncommon in pediatric rehabilitation, SLTs can now quantify dysarthria severity in children when using the severity scales of the p-RDA (after a training, if necessary) for a valid report on the child's ability to verbally communicate and to monitor motor speech development during treatment. The p-RDA is commercially available in the Netherlands for all SLTs working with childhood dysarthria.

## Acknowledgments

## Conflict of interest

The authors report no conflict of interest.

## References

[1]  Liégeois FJ, Morgan AT. Neural bases of childhood speech disorders: Lateralization and plasticity for speech functions during development. Neurosci Biobehav Rev. 2012; 36(1):439-58. doi: 10.1016/j.neubiorev.2011.07.011.

[2]  Dickinson HO, Parkinson KN, Ravens-Sieberer U, Schirripa G, Thyen U, Arnaud C, et al. Self-reported quality of life of 8-12-year-old children with cerebral palsy: a cross-sectional European study. Lancet. 2007;369(9580):2171-8. doi: 10.1016/s0140-6736(07)61013-7.

[3]  Fauconnier J, Dickinson HO, Beckung E, Marcelli M, McManus V, Michelsen SI, et al. Participation in life situations of 8–12 year old children with cerebral palsy: cross sectional European study. BMJ. 2009;338(232):1458. doi: 10.1136/bmj.b1458.

[4]  Parkes J, Hill N, Platt MJ, Donnelly C. Oromotor dysfunction and communication impairments in children with cerebral palsy: a register study: Oromotor and Communication Impairments in CP. Dev Med Child Neurol. 2010;52(12):1113-9. doi: 10.1111/j.1469-8749.2010.03765.x.

[5]  Cockerill H, Elbourne D, Allen E, Scrutton D, Will E, McNee A, et al. Speech, communication and use of augmentative communication in young people with cerebral palsy: The SH&PE population study: Communication and AAC use in bilateral cerebral palsy. Child Care Health Dev. 2014;40(2):149-57. doi: 10.1111/cch.12066.

[6]  Kooi-van Es M, Erasmus CE, de Swart BJ, Voet NB, van der Wees PJ, de Groot IJ, et al. Dysphagia and dysarthria in children with neuromuscular diseases, a prevalence study. J Neuromuscul Dis. 2020;7(3):287-95. doi: 10.3233/JND-190436.

[7]  Mary Watson R, Pennington L. Assessment and management of the communication difficulties of children with cerebral palsy: a UK survey of SLT practice. Int J Lang Commun Disord. 2015;50(2):241-59. doi: 10.1111/1460-6984.12138.

[8]  Knuijt S, Kalf JG, Gerven MH, Kocken JE, Kromhout LM, Goos HM, et al. Nederlandstalig dysartrieonderzoek volwassenen (NDO-V). Houten: Bohn Stafleu van Loghum; 2014.

[9]  Knuijt S, Kalf JG, van Engelen BG, de Swart BJ, Geurts AC. The Radboud dysarthria assessment: Development and clinimetric evaluation. Folia Phoniatr Logop. 2017;69(4): 143-53. doi: 10.1159/000484556.

[10]  Darley FL, Aronson AE, Brown JR. Motor Speech Disorders. Philadelphia: Saunders Company; 1975.

[11]  Enderby PM, John A. Therapy Outcome Measures (TOM): speech-language therapy. San Diego: Singular Publishing Group; 1997.

[12]  Hustad KC, Gorton K, Lee J. Classification of speech and language profiles in 4-year-old children with cerebral palsy: A prospective preliminary study. J Speech Lang Hear Res. 2010;53(6):1496-513. doi: 10.1044/1092-4388(2010/09-0176).

[13]  Morgan AT, Liégeois F. Re-thinking diagnostic classification of the dysarthrias: a developmental perspective. Folia Phoniatr Logop. 2010;62(3):120-6. doi: 10.1159/000287210.

[14]  van Haaften L, Diepeveen S, van den Engel-Hoek L, Jonker M, de Swart B, Maassen B. The psychometric evaluation of a speech production test battery for children: The reliability and validity of the Computer Articulation Instrument. J Speech Lang Hear Res. 2019;62(7):2141-70. doi: 10.1044/2018_JSLHR-S-18-0274.

[15]  McLeod S, Harrison LJ, McCormack J. The intelligibility in Context Scale: Validity and reliability of a subjective rating measure. J Speech Lang Hear Res. 2012;55(2):648-56. doi: 10.1044/1092-4388(2011/10-0130).

[16]  Hidecker MJ, Paneth N, Rosenbaum PL, Kent RD, Lillie J, Eulenberg JB, et al. Developing and validating the Communication Function Classification System for individuals with cerebral palsy: Developing a Communication Classification System. Dev Med Child Neurol. 2011;53(8):704-10. doi: 10.1111/j.1469-8749.2011.03996.x.

[17] Pennington L, Virella D, Mjøen T, da Graça Andrada M, Murray J, Colver A, et al. Development of The Viking Speech Scale to classify the speech of children with cerebral palsy. Res Dev Disabil. 2013;34(10):3202-10. doi: 10.1016/j.ridd.2013.06.035.

[18] Maas E, Jeukens-Visser M, Meester-Delver A, Beelen A. Interrater reliability of the capacity profile in children with neurodevelopmental disabilities. Arch Phys Med Rehabil. 2013;94(3):571-4. doi: 10.1016/j.apmr.2012.08.196.

[19] Eliasson A-C, Krumlinde-Sundholm L, Rösblad B, Beckung E, Arner M, Öhrvall A-M, et al. The Manual Ability Classification System (MACS) for children with cerebral palsy: scale development and evidence of validity and reliability. Dev Med Child Neurol. 2006;48(07):549. doi: 10.1017/S0012162206001162.

[20] Rosenbaum PL, Palisano RJ, Bartlett DJ, Galuppi BE, Russell DJ. Development of the Gross Motor Function Classification System for cerebral palsy. Dev Med Child Neurol. 2008;50(4):249-53. doi: 10.1111/j.1469-8749.2008.02045.x.

[21] Hidecker MJC, Ho NT, Dodge N, Hurvitz EA, Slaughter J, Workinger MS, et al. Inter-relationships of functional status in cerebral palsy: analyzing gross motor function, manual ability, and communication function classification systems in children: Relating the GMFCS, MACS, and CFCS. Dev Med Child Neurol. 2012;54(8):737-42. doi: 10.1111/j.1469-8749.2012.04312.x.

[22] Vignos PJ, Ge S, Archibald KC. Management of progressive muscular dystrophy in childhood. JAMA. 1963;184(2):89-96. doi:10.1001/jama.1963.03700150043007.

[23] Crary MA, Mann GD, Groher ME. Initial psychometric assessment of a functional oral intake scale for dysphagia in stroke patients. Arch Phys Med Rehabil. 2005;86(8):1516-20. doi: 10.1016/j.apmr.2004.11.049.

[24] Aaronson N, Alonso J, Burnam A, Lohr KN, Patrick DL, Perrin E, et al. Assessing health status and quality-of-life instruments: attributes and review criteria. Qual Life Res. 2002;11(3):193-205. doi: 10.1023/A:1015291021312.

[25] Terwee CB, Prinsen CA, Chiarotto A, Westerman MJ, Patrick DL, Alonso J, et al. COSMIN methodology for evaluating the content validity of patient-reported outcome measures: a Delphi study. Qual Life Res. 2018;27(5):1159-70. doi: 10.1007/s11136-018-1829-0.

[26] Forrest K. Diagnostic criteria of developmental apraxia of speech used by clinical speech-language pathologists. Am J Speech Lang Pathol. 2003;12(3):376-80. doi: 10.1044/1058-0360(2003/083).

[27] Shriberg LD, Kwiatkowski J. Developmental phonological disorders I: A clinical profile. J Speech Lang Hear Res. 1994;37(5):1100-26. doi: 10.1044/jshr.3705.1100.

[28] McCauley RJ, Strand EA. A review of standardized tests of nonverbal oral and speech motor performance in children. Am J Speech Lang Pathol. 2008;17(1):81-91. doi: 10.1044/1058-0360(2008/007).

[29] Kent RD, Kent JF, Rosenbek JC. Maximum performance tests of speech production. J Speech Hear Disord. 1987;52(4):367-387. doi:10.1044/jshd.5204.367.

[30] van Schie PE, Siebes RC, Dallmeijer AJ, Schuengel C, Smits DW, Gorter JW et al. Development of social functioning and communication in school-aged (5-9 years) children with cerebral palsy. Research in Developmental Disabilities. 2013;34(12):4485-4494. doi: 10.1016/j.ridd.2013.09.033.

[31] Ziegler W. Speech motor control is task-specific: Evidence from dysarthria and apraxia of speech. Aphasiology. 2003;17(1):33. doi: 10.1080/729254892.

[32] Hodge M, Wellman L. Management of children with dysarthria. In: Caruso E, Strand EA, editors. Clinical management of motor speech disorders in children. New York: Thieme; 1999.