UHASSELT

KNOWLEDGE IN ACTION

Faculteit Revalidatiewetenschappen

master in de revalidatiewetenschappen en de kinesitherapie

Masterthesis

Reliability and responsiveness study of a new assessment to evaluate sitting balance on the back of a horse in children with neurological disorders: The HippoTrunC

Lore Boers
Marie-Julie Bulen

Scriptie ingediend tot het behalen van de graad van master in de revalidatiewetenschappen en de kinesitherapie, afstudeerrichting revalidatiewetenschappen en kinesitherapie bij kinderen

PROMOTOR :
Prof. dr. Katrijn KLINGELS

BEGELEIDER :
Mevrouw Katrijne SEVERI

COPROMOTOR :
Prof. dr. Evi VERBECQUE

2022
2023

UHASSELT
KNOWLEDGE IN ACTION

www.uhasselt.be
Universiteit Hasselt
Campus Hasselt:
Martelarenlaan 42 | 3500 Hasselt
Campus Diepenbeek:
Agoralaan Gebouw D | 3590 Diepenbeek

UHASSELT

KNOWLEDGE IN ACTION

Faculteit Revalidatiewetenschappen

master in de revalidatiewetenschappen en de kinesitherapie

Masterthesis

Reliability and responsiveness study of a new assessment to evaluate sitting balance on the back of a horse in children with neurological disorders: The HippoTrunC

Lore Boers
Marie-Julie Bulen

Scriptie ingediend tot het behalen van de graad van master in de revalidatiewetenschappen en de kinesitherapie, afstudeerrichting revalidatiewetenschappen en kinesitherapie bij kinderen

PROMOTOR :
Prof. dr. Katrijn KLINGELS

COPROMOTOR :
Prof. dr. Evi VERBECQUE

BEGELEIDER :
Mevrouw Katrijne SEVERI

ACKNOWLEDGEMENT

This master's thesis marks the end of the five-year course in Rehabilitation Sciences and Physiotherapy at UHasselt. Within this program, we have gained knowledge, skills, and learned much about ourselves regarding self-development. We want to thank some people for the help and support they offered us in completing our education, and more specifically, writing the master's thesis.

First, we thank our supervisor and co-supervisor, Prof. dr. Katrijn Klingels and Prof. dr. Evi Verbecque. Both have provided constructive feedback over the two years, answered our questions, and encouraged us about our process and advice.

Furthermore, we would like to sincerely thank Sint-Gerardus and especially Katrijne Severi. Without her support, management, cooperation, and practical expertise, the study would have been impossible.

We would also like to thank each other for the smooth cooperation, the encouragement during more challenging moments, and the positive drive with which we infected each other.

Finally, we thank our parents and friends for their unconditional support, tips, and commitment. Your positive words and encouragement were much appreciated.

CONTEXT

This study on reliability and responsivity of the HippoTrunC is part of the research line in pediatric rehabilitation at REVAL. This project was conducted in collaboration with the school and care facility Sint-Gerardus, where all data was gathered. It was funded by "Het Innovatiefonds - VZW Stijn" and held at Diepenbeek, Belgium. Children with neurological problems, the majority of whom have impaired postural control, are the study's target population. A central format was used.

Children with neurological disorders struggle with various daily challenges, including maintaining balance while seated. Hippotherapy training approach has been shown to be a successful treatment method for addressing balance issues in children with neurological disorders (Novak & Honan, 2019). However, its impact on balance while seated on the back of a horse has yet to be substantially examined in the literature, which is understandable given the lack of a particular measurement method to evaluate sitting balance. Physiotherapists should assess sitting balance and analyse any significant changes after intervention, preferably while seated on the back of a horse, to provide the patient with the best possible patient-centred care. To fill this gap, the HippoTrunC was designed last year, which is specifically designed to measure sitting balance effects while seated on a horse after receiving hippotherapy (Hombroux & Van Dessel, 2022).

This study aims to determine the reliability and responsivity of the new measurement tool, the HippoTrunC.

The students and the master thesis's promotor developed the research question together. Supervisor K.S collected the data, which was followed by independent data analysis by both students. Consensus meetings were used to resolve disagreements.

1 ABSTRACT

Background: Postural instability is a common symptom in children with neurological disorders. Hippotherapy is considered an adequate therapy to positively influence postural control, balance and daily activities in this population. However, until today there is no standardised measurement tool that explicitly measures sitting balance to assess the effects of a hippotherapy intervention, which is why The HippoTrunC was developed.

Objectives: To investigate the reliability and responsiveness of the new measurement tool, the HippoTrunC, that measures the effect of sitting balance while seated on the back of the horse in children with neurological disorders.

Participants: The participants were 28 children with neurological disorders aged between 2 and 21 years. The mean age (SD) was 12.14 (4.4) years. All Gross Motor Function Classification Scale (GMFCS) levels were included.

Method: The HippoTrunC, consisting of 10 items on a 5-point scale, was administered three times while seated on the back of the horse. Test completions were filmed and scored afterwards. There was a one-week interval for test-retest reliability and a five-month interval for responsiveness. Intra-rater, inter-rater, and test-retest reliability were estimated at the scale level by calculating intraclass correlation coefficients (ICC). The weighted kappa was calculated to estimate intra-rater, inter-rater, and test-retest reliability at the item level. Responsiveness was assessed after five months of hippotherapy using the anchor-based approach (external criterion: the Goal Attainment Scale (GAS) and the Trunk Control Measurement Scale (TCMS)).

Results: The HippoTrunC showed excellent intra-, inter-, and test-retest reliability on the scale score (ICC= 0.976 - 0.998). Intra-, inter-, and test-retest reliability on item level varied between weighted kappas of 0.667 and 1.000. Results for responsiveness demonstrated that the mean change scores of the HippoTrunC were only significant for the GAS category 1 (scores 1 to 2 = progress larger than expected) (p= 0.003) and when the change score on the TCMS was equal or lower than the MCID of the TCMS (i.e. 6 points) (p= 0.003).

Conclusion: The results support the HippoTrunC as a reliable measurement tool. Responsiveness results show variable outcomes. Further research is needed to confirm these results and to assess other psychometric properties, such as validity.

Keywords: ‘sitting balance’, ‘neurological disorders’, ‘hippotherapy’.

2 INTRODUCTION

Postural control is a core feature of functioning, defined as the ability to control the centre of mass (COM) over the base of support (BOS) (Horak, 2009). To interact with the environment, make a coordinated movement, and maintain a posture such as sitting, it is required to have adequate postural control (Silkwood-Sherer et al., 2012). Sitting is an important posture for activities in daily life. Due to practice, experience, neural maturation, developing muscle strength, and changing biomechanical constraints, children learn to align their head, trunk and pelvis during sitting while maintaining their COM within the BOS (Inamdar et al., 2021). Sitting independently is an important developmental milestone in infancy in typically developing children but can be challenging for children with neurological disabilities (Inamdar et al., 2021).

Postural instability is a common symptom of paediatric movement disorders, which negatively impacts activities at school, at home, and during leisure activities (Silkwood-Sherer et al., 2012). Recent research has recognised that the global burden of disease is caused to a significant extent by neurological disorders in children leading to physical, cognitive, and psychosocial disabilities (Newton, 2018). Cerebral palsy (CP) is the most prevalent neurological cause of persisting motor function impairment, characterised by abnormalities of movement, muscle tone, and posture (O'Shea, 2008). CP is a heterogeneous clinical syndrome attributed to non-progressive disturbances usually caused by damage to the developing foetal or infant brain (Gulati & Sondhi, 2018). Clinical features of the movement disorder include spasticity (movement-dependent muscle stiffness), dyskinesia (uncontrollable movements), ataxia (poor coordination), and poor trunk control (Vitrikas et al., 2020). Poor trunk control affects sitting and walking. As the relationship between trunk control in sitting and during walking has been established, it is necessary to prioritise trunk control in sitting during treatment (Sæther et al., 2015).

In the past decades, top-down interventions have been expanding, which focus on improving the degree of independence during daily activities and participation. In their systematic review, Novak and colleagues have shown the effectiveness and evidence of several therapies, among which hippotherapy is highly recommended to improve postural balance in children (Novak & Honan, 2019).

Given the importance of hippotherapy on balance in different populations, clinicians and therapists must employ a standardised measurement tool. There are several measurement tools for children with CP available to measure sitting balance that have a high level of evidence, including the Trunk Control Measurement Scale (TCMS), the Level of Sitting Scale (LSS), the Segmental Assessment of Trunk Control (SATCo) (Saether et al., 2013), the sitting subscale of the Pediatric Reach Test (PRT) and the Sitting Assessment for Children with Neuromotor Dysfunction (SACND) (Bañas & Gorgon, 2014).

To the best of our knowledge, no standardised measurement tool is available in the literature that precisely measures sitting balance while seated on the back of the horse, which is why the HippoTrunC was developed. This test aims to detect changes in sitting balance in children with neurological disorders after hippotherapy interventions. The HippoTrunC is administered while seated on the back of the horse. It contains a total of 10 items divided into three domains of sitting balance: static, active and reactive control. A score (1–5) is assigned based on the degree of the head and trunk alignment and the level of anticipation, the use of aids and the degree of manual support while sitting on the back of the horse during specific test conditions. Preliminary investigations in a small sample of children with neurological disorders (n=12) have shown that the test is feasible and reliable. Due to the lower intra-rater reliability at the item level, several items still needed to be fine-tuned to draw general conclusions (Hombroux & Van Dessel, 2022). The HippoTrunC was adjusted based on a previous pilot study to obtain more consistency. Psychometric properties need to be investigated for this final version of the HippoTrunC.

This thesis presents the hypothesis that the HippoTrunC is reliable on inter- intra- and test-retest reliability and responsive based on an anchor-based method of the GAS (changes relevant to the patient's needs) and TCMS (trunk control while seated). This study is relevant as the HippoTrunC provides added importance to the measurement of sitting balance in further research and for use in clinical settings in children with neurological disorders.

Therefore, this study aims to investigate the reliability and responsiveness of the HippoTrunC.

3 METHODE

3.1 Study design

A prospective, observational, longitudinal study was conducted to establish the different types of reliability (intra-rater, inter-rater, and test-retest) and responsiveness of a newly composed measurement scale. Measurements took place at baseline in October 2022 with a one-week follow-up for determining test-retest reliability and five months later to establish responsiveness.

3.2 Participants

Thirty-five children from school and healthcare facility "Sint-Gerardus" in Diepenbeek, Belgium were invited to participate in this study. To be included in this study, children had to be aged between 2 and 21 years, able to communicate (none) verbally, have a neurological disorder and have received at least one hippotherapy session before the start of the study. All Gross Motor Function Classification System (GMFCS) levels were included in this study. Participants in this study were excluded if they were unable to comprehend or follow short, simple instructions.

All possible participants and their parents/guardians were contacted, and consent forms were signed before they could participate in the study. This project was approved on the 17th of December 2021 by the University of Hasselt Medical Ethics Committee (B1152021000028).

3.3 Outcome measures

3.3.1 Primary outcome measure

Measurement tool "Hippotherapy Trunk Control Scale"

The 'Hippotherapy Trunk Control Scale' (HippoTrunC), is a measurement tool specifically designed to assess sitting balance, and head and trunk control while sitting on the back of a horse. The HippoTrunC includes 10 items scored with a 5-point rating scale (1: unable to perform, 5: able to complete the task as instructed without difficulty). The total scores range from 6 – 50 because a minimum score of 10 can be obtained on all items together, but a maximum of four points can be deducted from this score when support is needed during the test (stirrups, corset and/or duo-sit). The HippoTrunC was composed by deriving items from

three existing reliable and valid assessment tools used to evaluate sitting balance: the Sitting Assessment Scale (SAS), SATCo and Gross Motor Function Measure (GMFM) (Adair et al., 2012; Butler et al., 2010). A previous study has shown that the HippoTrunC is user-friendly and feasible to apply in this specific context, but adjustments have been made to make the scale more reliable (Hombroux & Van Dessel, 2022). The instructions for all items were modified so that the child performs the items as predetermined, and the scale became a 5-point scale instead of a 4-point scale. This allows a more detailed differentiation during scoring and better discrimination. The scale allows assistive devices (corset, duo-sit, and stirrups), but also corrects for it by subtracting one or two points from the total score for each level of support. Furthermore, the course set-up is also changed to make the test more efficient. The four GoPro's were relocated to different places so that the video recordings could be scored more accurately.

Appendix Table 1 & 2 provides an overview of the adjustments and composition of the HippoTrunC.

3.3.2 Secondary outcome measures

The anchor-based approach was used to evaluate if the changes measured with the HippoTrunC corresponded to changes reported by the treating therapist. This made it possible to verify whether observed changes were relevant to the patient's needs (GAS as an anchor) or related to the trunk control while seated (TCMS as an anchor).

Goal Attainment Scaling (GAS)

Goal Attainment Scaling (GAS) is an individualised evaluation method. It is scored on an ordinal 5-point scale that captures a person's individual treatment goal and reassesses these goals after treatment by mapping the degree to which they have been achieved. The GAS has good inter-rater reliability and validity, and it presents excellent sensitivity to change in different populations (cognitive, paediatric, and neurological target groups) (Krasny-Pacini et al., 2013). Two goals were defined at the function level based on activities in daily life and scaled before the first test session. The expert in the field, K.S., who gives hippotherapy, defined goals for each participant in the context of the trunk, head control and sitting balance, thereby individualising the goals.

Appendix Figure 1 provides an example of the GAS.

Trunk control measurement scale (TCMS)

The trunk control measurement scale (TCMS) consists of static and dynamic sitting balance, where dynamic sitting balance is divided into selective motion control and dynamic reaching (Lopez-Ruiz et al., 2023). The TCMS is a 15-item scale scored on an ordinal 2-, 3-, or 4-point rating scale. The total score of the TCMS ranges from 0 (low performance) to 58 (high performance). Measurement properties such as reliability (ICC vary for inter-rater and test-retest reliability between 0.91 and 0.99) and validity (the correlations range between 0.57 and 0.75 ($p<0.003$)) are good in 8- to 15-year-old children with CP (Pham et al., 2016).

3.4 Procedure

The procedure is divided into several steps over time, summarised in ***Figure 1.***

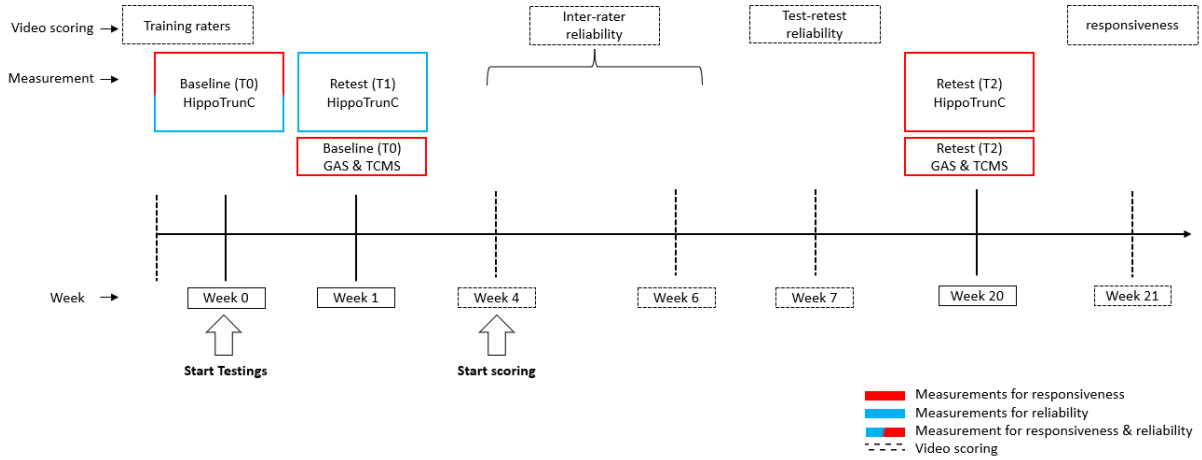


Figure 1: Summary of the Procedure

3.4.1 Evaluation

The raters (master students LB and MJB) received training in using the HippoTrunC. Twelve videos were to be scored in three blocks of four videos, and feedback on scoring was provided after completing each block. At baseline, therapists provided information on the child's sex, age, GMFCS level, diagnosis, and assistive devices used.

The participants were tested for the first time in October 2022. This test took place at 'Sint-Gerardus' in Diepenbeek during the hippotherapy sessions. The test session of each child was recorded by a setup with four GoPro's at different angles (HERO8). ***Appendix Figure 2*** provides an overview of the setup. Each test was afterwards scored based on video recordings by two independent raters (LB and MJB), and in case of disagreement, a consensus meeting was held. If the assessors could not reach a consensus, the opinion of a third rater

was decisive (KS). The physiotherapist at school, who gives hippotherapy, administered the HippoTrunC, GAS and TCMS from the children.

3.4.2 Assessment for reliability and responsiveness

As shown in ***Figure 1***, the participants were reassessed one week after the baseline measure to evaluate test-retest reliability and after five months for responsiveness.

The first GAS administration was conducted one week later as the second test session of the HippoTrunC. After a therapy period of five months, participants were invited to participate in a third testing session, during which time the HippoTrunC was administered again, along with the GAS and the TCMS. The change in HippoTrunC was then correlated to the level of change in their functional goal and the degree of improvement in trunk control.

3.4.2.1. Reliability

Reliability refers to the extent to which a measurement tool will produce similar results in different circumstances, assuming nothing else has changed (Roberts & Priest, 2006). Three different types of reliability were examined. Intra-rater reliability measures how consistent one assessor is at measuring the same data, and inter-rater reliability measures how consistent independent assessors are at measuring the same data (McHugh, 2012). Test-retest reliability reflects the consistency of the measurement tool and can be measured by having the same group of participants take the same test at two separate time points (Story & Tait, 2019).

Two researchers independently scored with at least 14 days in between for intra-reliability. For inter-reliability, both researchers scored simultaneously. To measure test-retest reliability, participants were tested a second time with a one-week interval, and both researchers scored simultaneously.

3.4.2.2 Responsiveness

Responsiveness refers to the ability of an outcome measure to detect change over time in the construct to be measured and thus refers to the validity of a change score (Mokkink et al., 2010). There was a baseline measurement at the start of the school year. After a follow-up of five months, the second measurement was conducted (***Figure 1***).

3.5 Statistical analyses

3.5.1 Reliability

SPSS Statistics ("SPSS Statistics for Windows," 2021) was used for statistical analysis. Intra-rater, inter-rater, and test-retest reliability of the ***scale score*** were determined with ICCs. A 2-way random effects model was used for inter-rater reliability and a 2-way mixed model for intra-rater and test-retest reliability. The ICC ranges from 0-1 with the following interpretation: < 0.5: poor reliability; between 0.5- 0.75: moderate reliability; between 0.75-0.9 good reliability; > 0.90: excellent reliability (Koo & Li, 2016). Differences between scores can be explained by disagreements between different assessors (Hallgren, 2012). In addition, for test-retest reliability at the scale level, the Standard Error of Measurement (SEM; formula: SEM =SD √ (1− ICC), and the Smallest Detectable Change (SDC); formula: SDC = 1.96*√2*(SEM)) was determined.

Intra-rater, inter-rater, and test-retest reliability at the ***item level*** were evaluated using the quadratic weighted kappa with a 95% confidence interval. This quadratic weighted kappa coefficient is used to compare the variability between pairs of items with the total variability (Vanbelle, 2016). The weighted kappa ranges from 0-1 with the following interpretation: values ≤0 as having no agreement, 0.01-0.20 as none to slight agreement, 0.21-0.40 as fair, 0.41- 0.60 as moderate, 0.61-0.80 as substantial, and 0.81-1.00 as near-perfect agreement (McHugh, 2012; Vanbelle, 2016).

3.4.2 Responsiveness

Responsiveness was evaluated at the scale level. The minimum clinically important difference (MCID) was used to determine the responsiveness of the HippoTrunC results. The MCID for the HippoTrunC using the anchor-based approach was estimated following two steps. First, for both anchors (GAS and TCMS), a baseline measurement and a follow-up measurement (5 months after baseline) were conducted to capture the degree of change. After five months of therapy, change was assessed by the therapist. Between each GAS category in which change was represented, HippoTrunC scores were assessed for significant differences, for which the mean difference between baseline and follow-up was analysed. The goals were plotted on a 5-point scale, in which the mean change of participants was defined as -2 (baseline/not improved), -1 (minimal change), 0 (expected change), and +1/+2 (more than expected

change). The change scores were divided into three groups: -1 (scores from -2 to -1 = status quo), 0 (score expected change), and 1 (scores from 1 to 2 = progress larger than expected). The same procedure was used for the TCMS as an anchor. Lopez-Ruiz et al., 2023 showed an MCID of 5.15 points at a 95% confidence interval for the TCMS. A significant change was considered from an MCID of 6 points difference in pre- and post-score. The MCID was used as a cut-off value to classify the children: 0 (no change, < 6 points) and 1 (significant change, ≥ 6 points).

Second, a Wilcoxon signed-rank test was used to determine the significance (< 0.05) between baseline and follow-up scores for each GAS category and for each TCMS category. The mean changes in the scores of the HippoTrunC were calculated from baseline to follow-up for each category of both anchors, providing information about the spread of the mean difference within each category, revealing significant differences. This spread was plotted in graphs.

Additionally, Spearman rank correlation coefficients and scatter plots were used to explore the relationship between TCMS and HippoTrunC (TCMS baseline scores vs HippoTrunc baseline scores and TCMS differentiation vs HippoTrunC differentiation). The correlations range from zero to one with the following interpretation: 0.00- 0.25: little or no relationship; 0.25-0.50: fair relationship; 0.50-0.75: moderate to good relationship; 0.75-1.00: good to excellent relationship (Portney & Watkins, 2009).

4 RESULTS

4.1 characteristics of the participants

Out of 35 invited participants, 30 children gave their consent. Throughout the study, two participants dropped out during the first week, leaving 28 children eligible for data analysis of the reliability results. Another four participants dropped out after the five-month follow-up period, leaving 24 children eligible for analysis. Reasons for dropouts were medical reasons, school trips, and absence from school because of social-emotional problems. Due to the fact that two participants only dropped out in March, reliability results were based on 28 participants, but responsiveness results were based on 24 participants. The mean age (SD) was 12.1 (4.4) years, and a sex distribution of 8 females and 20 males participants. Eleven children were classified with GMFCS level IV. Twenty-one children were classified with CP, and other neurological disorders like spina bifida and Duchene were also included and classified on the GMFCS. ***Table 1*** provides a description of the participants with respect to age, sex, diagnosis, and degree of disability.

Table 1

Characteristics of the Participants

	Age(years)	Sex	Diagnosis	Degree of disability (GMFCS)				
				I	II	III	IV	V
1	12.8	M	CP (bilateral)	x				
2	9.0	M	CP (bilateral)				x	
3	12.8	M	CP (bilateral)				x	
4	4.9	F	CP (bilateral)		x			
5	10.3	F	CP (bilateral)					x
6	14.4	F	CP (bilateral)					x
7	9.2	F	CP (bilateral)			x		
8	9.5	M	CP (bilateral)					
9	17.0	M	CP (bilateral)				x	
10	4.6	M	Spina bifida				x	
11	15.8	F	CP (bilateral)			x		
12	15.4	M	CP (bilateral)				x	
13	7.6	M	CP (bilateral)					x
14	17.10	M	CP (bilateral)					x
15	8.3	F	Sturge-Weber syndrome					x
16	17.0	F	CP (bilateral)				x	
17	12.8	M	DMD			x		
18	8.9	M	CP (unilateral)		x			
19	18.12	M	CP (bilateral)					x
20	15.9	M	CP (bilateral)				x	
21	7.7	M	DIPG			x		
22	17.2	M	CP (bilateral)				x	
23	15.3	M	Cris-du-chat			x		
24	7.5	M	CP (unilateral)		x			
25	14.3	M	CP (bilateral)				x	
26	6.2	M	Supratentoriële hydrocephalus				x	
27	10.3	M	CP (bilateral)				x	
28	20.11	F	Goltz syndrome	x				

Abbreviations. F = Female; M = Male; GMFCS = Gross Motor Function Classification Scale; CP = Cerebral Palsy; DMD = Duchenne muscular dystrophy; DIPG = Diffuse intrinsic pontine glioma;
Explanation: Sturge-Weber syndrome = Facial capillary malformation combined with ocular and neurological problems; Goltz syndrome = An inherited disorder with an increased risk of cancer and tumour formation

4.2 Reliability

4.2.1 Intra-rater reliability

For the scale level, the agreement between the scale scores assigned by rater 1 (ICC= 0.998, 95% CI= [0.974; 0.994]) and rater 2 (ICC= 0.995, 95% CI= [0.989; 0,998]) was excellent.

For the item level, agreement by rater 1 varied between 0.667 (item 6) and 1.000 (item 10). All items showed near-perfect agreement, only one item (item 6) showed substantial agreement. Agreement by rater 2 varied between 0.725 (item 1) and 1.000 (items 4, 9, and 10). One item showed substantial agreement (item 1). All the other items showed near-perfect agreement between the two measurement times.

Table 2 provides an overview of the weighted Kappa's for each item and a 95% CI.

Table 2

Intra-Rater Reliability – Item Level

Item		Rater	Weighted Kappa	95% confidence interval	
				Lower bound	Upper bound
1	Head control: static	1	0.826	0.688	0.964
		2	0.725	0.538	0.911
2	Sitting balance (A): static	1	0.973	0.945	1.002
		2	0.899	0.773	1.025
3	Sitting balance (A): forward reaching	1	0.989	0.967	1.010
		2	0.965	0.893	1.036
4	Sitting balance (A): lateral reaching	1	0.962	0.931	0.994
		2	1.000	1.000	1.000
5	Sitting balance: active	1	0.916	0.814	1.019
		2	0.977	0.945	1.010
6	Head control: active	1	0.667	0.377	0.956
		2	0.948	0.863	1.033
7	Sitting balance (A): acceleration	1	0.968	0.934	1.001
		2	0.986	0.959	1.013
8	Sitting balance (A): slalom	1	0.991	0.975	1.008
		2	0.968	0.931	1.005
9	Sitting balance (R): slalom	1	0.988	0.965	1.010
		2	1.000	1.000	1.000
10	Sitting balance (R): start-stop	1	1.000	1.000	1.000
		2	1.000	1.000	1.000

Abbreviation. A=active, R=reactive

4.2.2 Inter-rater reliability

For the scale level, intraclass correlation coefficients demonstrated excellent inter-rater reliability for the scale score (ICC= 0.994, 95% CI= [0.987; 0.997]).

For the item level, agreement varied between 0.742 (item 6) and 0.977 (item 4). One item showed a substantial agreement (item 6). All the other items showed an excellent agreement between both assessors. ***Table 3*** provides an overview of the weighted Kappa's for each item and a 95% CI.

Table 3

Inter-Rater Reliability – Item Level

Item		Weighted Kappa	95% confidence interval	
			Lower bound	Upper bound
1	Head control: static	0.836	0.717	0.954
2	Sitting balance (A): static	0.948	0. 910	0.986
3	Sitting balance (A): forward reaching	0.951	0.916	0.986
4	Sitting balance (A): lateral reaching	0.977	0.739	1.014
5	Sitting balance: active	0.881	0.782	0.980
6	Head control: active	0.742	0.528	0.956
7	Sitting balance (A): acceleration	0.936	0.881	0.990
8	Sitting balance (A): slalom	0.931	0.851	1.011
9	Sitting balance (R): slalom	0.907	0.793	1.021
10	Sitting balance (R): start-stop	0.913	0.804	1.023)

Abbreviation. A=active, R=reactive

4.2.3 Test-retest reliability

For the scale level, the test-retest reliability value for the total score of the HippoTrunC was excellent (ICC=0,976; 95% CI= [0.949; 0.989]) with an SEM of 1.708 and an SDC of 4.735. Seven of the 10 items showed a near-perfect agreement for the item level, and the remaining 3 items (items 8, 9, and 10) showed a substantial agreement. The variation ranged from 0.722 (item 9) to 0.904 (item 2). The test-retest reliability for the item level is summarised in ***Table 4.***

Table 4

Test-Retest Reliability – Item Level

Item		Weighted Kappa	95% confidence interval	
			Lower bound	Upper bound
1	Head control: static	0.819	0.671	0.968
2	Sitting balance (A): static	0.904	0. 828	0.981
3	Sitting balance (A): forward reaching	0.888	0.787	0.990
4	Sitting balance (A): lateral reaching	0.873	0.760	0.985
5	Sitting balance: active	0.862	0.760	0.963
6	Head control: active	0.832	0.696	0.968
7	Sitting balance (A): acceleration	0.883	0.798	0.968
8	Sitting balance (A): slalom	0.768	0.614	0.922
9	Sitting balance (R): slalom	0.722	0.545	0.899
10	Sitting balance (R): start-stop	0.763	0.605	0.921

Abbreviations. A= active, R= reactive

4.3 Responsiveness

Five participants scored higher than the MCID (4.7 points) of the HippoTrunC. Data analysis for the GAS as anchor indicated that the mean change scores of the HippoTrunC were only significant for the GAS category 1 (scores from +1 to +2 = progress larger than expected) in both the first goal (p = 0.003) and the second goal (p= 0.003) of the GAS. After a five-month follow-up. This indicates that participants who improved by 1 or 2 points on the GAS after the five-month follow-up significantly improved on the HippoTrunC. These significant results were not found in the GAS category 0 (score expected change and GAS category -1 (scores from -2 to -1 = status quo). The results are shown in ***Figures 2 & 3***.

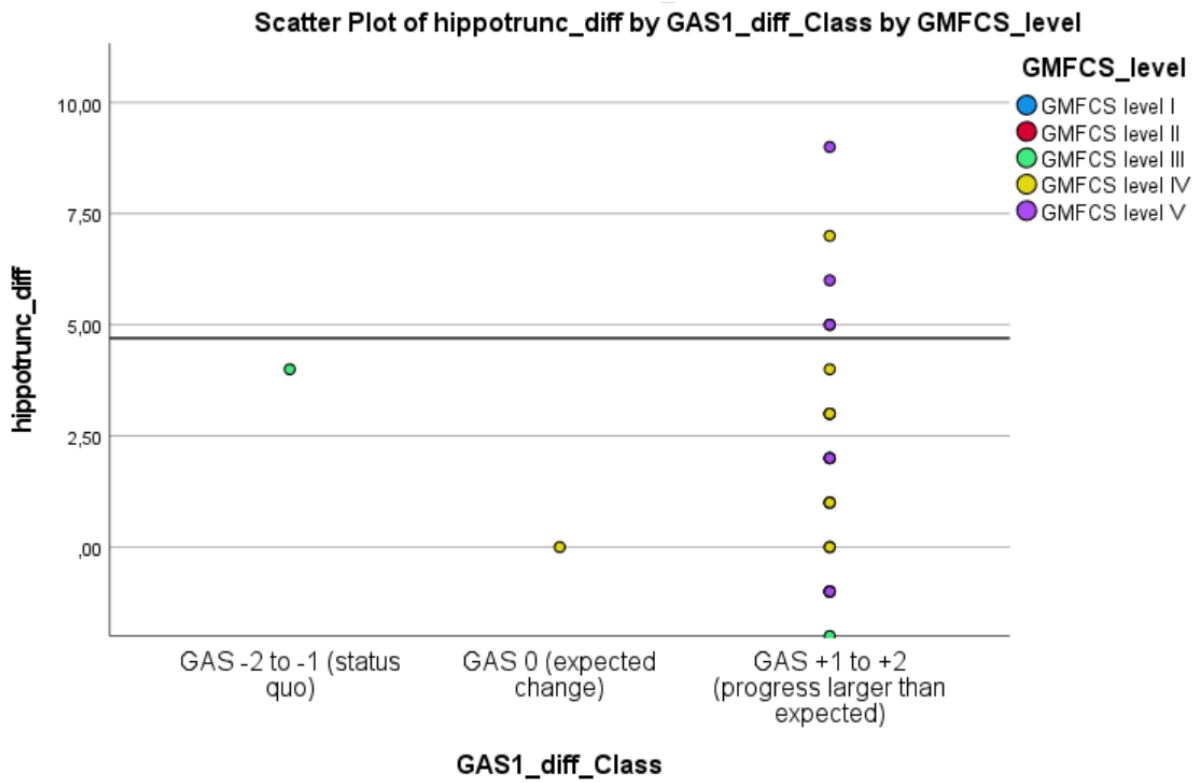


Figure 2. Scatter Plot of Goal 1 of the GAS

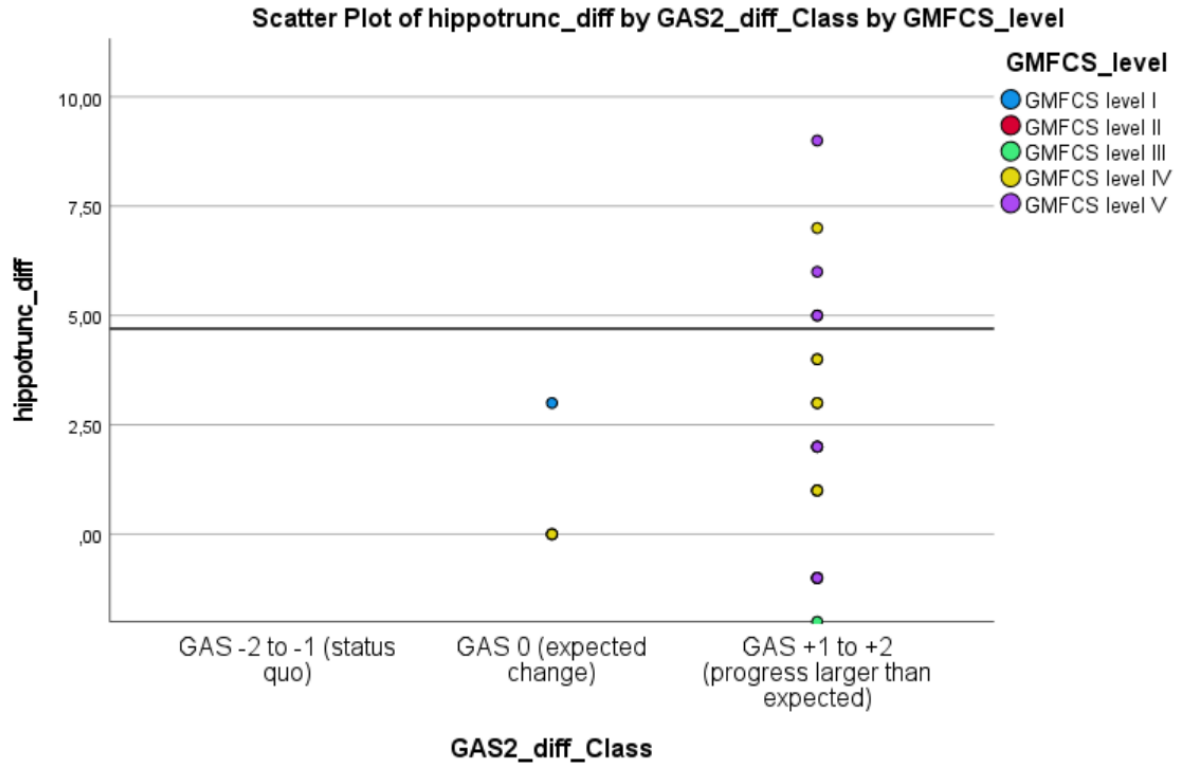


Figure 3. Scatter Plot of Goal 2 of the GAS

The study results using TCMS as an anchor showed that the change score of the HippoTrunC was significant only when the change score on the TCMS was equal to or lower than the MCID of the TCMS (i.e. 6 points) (p= 0.003), with nineteen participants scoring equal or lower than the MCID. When the change score on the TCMS was more than the MCID, no significant improvement was seen in the HippoTrunC score. The results are shown in ***Figure 4***.

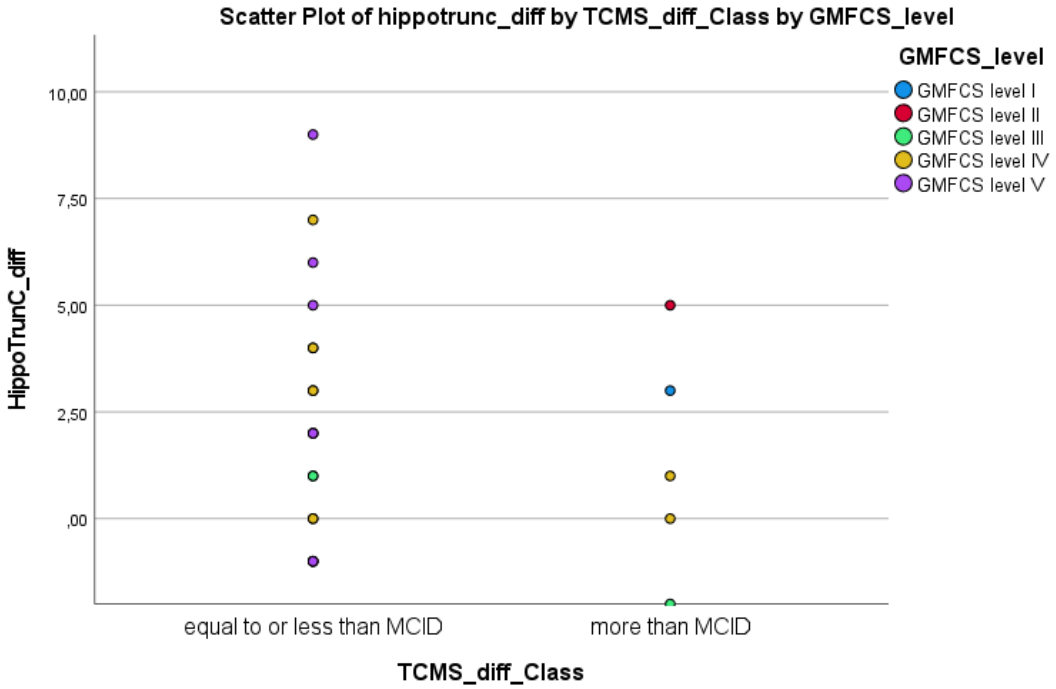


Figure 4. Scatter Plot TCMS

Spearman rank correlation coefficients between TCMS baseline scores and HippoTrunC baseline scores showed an excellent correlation (p= 0.885). The correlation between TCMS differentiation (TCMS post - TCMS pre) and HippoTrunC differentiation (HippoTrunC post - pre) showed little or no relationship (p = 0.053). ***Appendix Figure 3 & 4*** provides an overview of the scatter plots for the correlations.

5 DISCUSSION

The aim of this study was to investigate whether the HippoTrunC is reliable and responsive. The relevance of this study consists of the potential use of this measurement scale in a clinical setting and in intervention studies to improve sitting balance. This study showed that the HippoTrunc is a reliable outcome measure for evaluating sitting balance while seated on the back of a horse in children with neurological disorders. The HippoTrunC showed excellent intra-rater, inter-rater and test-retest reliability for the scale score. Most items also showed near-perfect agreement for intra-rater, inter-rater, and test-retest reliability at the item level. The present study is the first to examine the responsiveness of the HippoTrunc, which showed to be variable depending on the anchor.

Reflection of findings related to the research question

Last year's pilot study also noted the results of inter- and intra-rater reliability on the scale level and intra-rater on the item level (Hombroux & Van Dessel, 2022). For inter-rater reliability at the item level, lower weighted kappa values were found in the previous study (Hombroux & Van Dessel, 2022). These improved outcomes may be due to fine-tuning the items open for interpretation, changing to a 5-point scale and relocating the GoPro's. This means that the adjustments made are beneficial for inter-rater reliability on the item level, where now the majority of the items showed near-perfect reliability, and only one item showed substantial agreement, i.e. item 6, which measures active head control. Test-retest reliability was not previously examined for the item and scale level. For the item level, 3 of the 10 items showed substantial agreement; these items evaluated active and reactive sitting balance. Not obtaining excellent agreement may be due to interpreting these items when the trunk is in motion, and it isn't easy to decide when the trunk has moved inadequately.

Findings about reliability in children with neurological disorders in previous studies suggest the potential applicability of several sitting balance measurement tools. Excellent inter-rater reliability outcomes were reported for the sitting subscale of the PRT, SATco, and TCMS. Inter-rater reliability estimates for the SACND ranged from poor to excellent. The SACND and TCMS test-retest reliability values were excellent; for the sitting subscale of the PRT, they ranged from poor to excellent. The SATCo showed excellent intra-rater reliability (Bañas & Gorgon,

2014). Therewithal, Heyrman et al., 2011 showed excellent reliability of the TCMS on the total score and subscales for inter-rater reliability and test–retest reliability.

For responsiveness, it is not possible to draw unequivocal conclusions, given the variable results. When the GAS was used as an anchor, a significant improvement on the HippoTrunC was visible when the GAS improved. The physiotherapist at school defined the GAS goals at the function level, e.g. "keeping the trunk straight or trunk response to abrupt movements". These goals are important for performing activities of daily living, such as independent reaching while sitting. Although participants may not perceive a significant improvement on the GAS goals as a significant change or a real change at the activity level. For further research, it is recommended to set goals at the activity or participation level to make it more apparent what the clinical relevance is in daily life for the participants. Carey et al., 2016 did a similar study with children with CP, examining the mean change score (baseline - follow-up) of the Time Up and Go (TUG), a test for postural control, balance, and functional mobility using the GAS as an anchor. The patients, themselves or their families, selected two meaningful goals for the GAS: one at the activity and one at the participation level. The change scores on the GAS were divided into four groups: -2 (no change), -1 (minimal change), 0 (moderate change), and 1 or 2 (large change). For most participants, the mean change in the TUG score was significant in the minimal change category (-1) on the GAS. These results showed different results from our study, where only the mean change scores of the HippoTrunC were significant for the GAS category 1 (scores from +1 to +2 = progress larger than expected). A possible hypothesis could be that the goals were made by patients of the family instead of a caregiver and were located at the activity of participation level. Given the limited sample size, this study is divided into three categories of the GAS. Ideally, future goals are divided among five categories of the GAS when the sample size is larger and are localised at the activity or participation level.

An earlier responsiveness study shows a moderate correlation (p= 0.54) between TCMS and hippotherapy training in children and adolescents with CP (Pham et al., 2016). In this study, the TCMS and the HippoTrunC baseline showed an excellent correlation, indicating that the scales measure the same; they both evaluate static and dynamic sitting balance. However, the relationship between the difference after the five-month follow-up showed no significant

correlation, which aligns with our results. Compared to the change on the TCMS, the change on the HippoTrunC was the other way around. Children who did not change significantly on the TCMS did on the HippoTrunC and vice versa. Multiple hypotheses about this can be made. The first hypothesis states that it is easier to go from a score of 1 or 2 to a score of 3 than from a score of 4 to 5 on the HippoTrunC. The degree of support and/or the extent of trunk movement and/or the distance of trunk/head righting is decisive in several items. More specifically, to go from a score of 2 to 3, the participant must change from moderate to mild support and/or from large to moderate trunk movements and/or independent righting for a few strides to righting for half the length. A score of 5 implies perfect execution without compensations, making the difference between scores 4 and 5 large. The second hypothesis is that the TCMS allows more progression for better-performing children; for weaker children, it is difficult to progress. More in detail, it can be seen that children received a total score of 0 if they scored 0 on item 1 (patient falls or can only maintain upright sitting with double arm support). In this case, the participant is not allowed to perform the next TCMS resulting in very low scores. Any minor improvement, e.g. improvement from score 0 to 1 on item 1, can therefore have a major impact on the overall score. Heyrman et al., 2011 show similar results, concluding that the TCMS is suitable for children who can maintain a sitting position independently and is, therefore, more likely to focus on GMFCS levels I to III, for GMFCS levels V or IV, they suggest other measurement instruments, such as the SATCo. Based on these hypotheses, it could be argued that the HippoTrunC is more sensitive to minor changes, making the HippoTrunC suitable for children with a lower GMFCS level, but this should be investigated further with a larger sample size.

Strengths and limitations of the study

A notable strength of this study was the sample size of 28 children with different neurological disorders. Secondly, The HippoTrunC requires only little equipment and can be performed within 15 minutes, which makes it a practical tool. Thirdly, the video recordings from different angles accurately showed the children's performances which can be watched multiple times. Furthermore, there were two independent raters for administering the HippoTrunC, reducing the risk of observer bias. The interpretation of the results might be influenced due to methodological limitations, such as no equal distribution of participants regarding sex and GMFCS levels. A higher percentage of boys and children with level GMFCS IV participated.

Consequently, generalising the results should be done with caution. Secondly, the measurements of the HippoTrunC were done with horses, so you cannot control everything as a tester/therapist. Difficulties here were when the horse had to stand still for a certain period. Some items were open to interpretation because of the description of the items. For example, in items 1 and 6, it was difficult to determine what an upright position is while seated on the back of a horse since no clinical examination was performed beforehand. Apart from the open interpretation, there was a possibility of a learning effect for item 10. Some children know after doing the test once or twice that the horse stops abruptly twice at this item, which allows them to prepare for this. Furthermore, due to a lack of understanding or motivation, some participants required more explanation or more than two attempts. Finally, two limitations emerged when the test was administered in duo-sitting. It was not easy to score through the video recordings how much support was given, and two or three people must be involved depending on solo or duo sitting when taking the test.

Implications for clinical practice and further research

The HippoTrunC enables, the assessment of strengths and weaknesses in children with neurological disorders while seated on the back of a horse. By analysing the scoring, it is possible to differentiate between static versus dynamic balance and trunk versus head control. The HippoTrunC makes it possible to measure evolution over time and register objective data. Furthermore, the HippoTrunC has a small SDC, which supports its use to objectify relevant changes in long-term studies. The measurement tool needs to be further tested in clinical settings to confirm the promising results about the HippoTrunC's reliability and responsiveness. Additionally, a larger sample is needed with an equal division in the variation of different neurological disorders, gender, and level of GMFCS. The HippoTrunC should be expanded across different places or countries to test cross-cultural validation. For this reason, the test should also be released in other languages. There are still more psychometric aspects to be determined, such as validity, in addition to reliability and responsiveness. For instance, to look into a potential floor or ceiling impact further.

6 CONCLUSION

Our results support the reliability of the HippoTrunC in children with neurological disorders. Variable results are shown for responsiveness. Current results seem to support that the HippoTrunc is sensitive to detect large changes in personal goals. Compared to the TCMS, the HippoTrunC only detects differences in children who score weakly on the TCMS. This makes the HippoTrunc seem more suitable for children who perform weakly than for stronger-performing children. This may be due to the scale's sensitivity for more severely affected children, but may also be due to other issues that need further exploration in follow-up studies. To draw general conclusions about responsiveness, it is necessary to have a larger number of participants and goals on GAS need to be on activity level. Further research is needed to improve responsiveness and to investigate other psychometric properties like validity.

7 REFERENCES

Adair, B., Said, C. M., Rodda, J., & Morris, M. E. (2012). Psychometric properties of functional mobility tools in hereditary spastic paraplegia and other childhood neurological conditions. *Dev Med Child Neurol*, *54*(7), 596-605. https://doi.org/10.1111/j.1469-8749.2012.04284.x

Bañas, B. B., & Gorgon, E. J. (2014). Clinimetric properties of sitting balance measures for children with cerebral palsy: a systematic review. *Phys Occup Ther Pediatr*, *34*(3), 313-334. https://doi.org/10.3109/01942638.2014.881952

Butler, P. B., Saavedra, S., Sofranac, M., Jarvis, S. E., & Woollacott, M. H. (2010). Refinement, reliability, and validity of the segmental assessment of trunk control. *Pediatr Phys Ther*, *22*(3), 246-257. https://doi.org/10.1097/PEP.0b013e3181e69490

Carey, H., Martin, K., Combs-Miller, S., & Heathcock, J. C. (2016). Reliability and Responsiveness of the Timed Up and Go Test in Children With Cerebral Palsy. *Pediatr Phys Ther*, *28*(4), 401-408. https://doi.org/10.1097/PEP.0000000000000301

Gulati, S., & Sondhi, V. (2018). Cerebral Palsy: An Overview. *Indian J Pediatr*, *85*(11), 1006-1016. https://doi.org/10.1007/s12098-017-2475-1

Hallgren, K. A. (2012). Computing Inter-Rater Reliability for Observational Data: An Overview and Tutorial. *Tutor Quant Methods Psychol*, *8*(1), 23-34. https://doi.org/10.20982/tqmp.08.1.p023

Heyrman, L., Molenaers, G., Desloovere, K., Verheyden, G., De Cat, J., Monbaliu, E., & Feys, H. (2011). A clinical tool to measure trunk control in children with cerebral palsy: the Trunk Control Measurement Scale. *Res Dev Disabil*, *32*(6), 2624-2635. https://doi.org/10.1016/j.ridd.2011.06.012

Horak, F. B. (2009). Postural Control. In *Encyclopedia of Neuroscience* (pp. 3212-3219). https://doi.org/10.1007/978-3-540-29678-2_4708

Inamdar, K., Molinini, R. M., Panibatla, S. T., Chow, J. C., & Dusing, S. C. (2021). Physical therapy interventions to improve sitting ability in children with or at-risk for cerebral palsy: a systematic review and meta-analysis. *Dev Med Child Neurol*, *63*(4), 396-406. https://doi.org/10.1111/dmcn.14772

Koo, T. K., & Li, M. Y. (2016). A Guideline of Selecting and Reporting Intraclass Correlation Coefficients for Reliability Research. *J Chiropr Med*, *15*(2), 155-163. https://doi.org/10.1016/j.jcm.2016.02.012

Krasny-Pacini, A., Hiebel, J., Pauly, F., Godon, S., & Chevignard, M. (2013). Goal attainment scaling in rehabilitation: a literature-based update. *Ann Phys Rehabil Med*, *56*(3), 212-230. https://doi.org/10.1016/j.rehab.2013.02.002

Lopez-Ruiz, J., Estrada-Barranco, C., Martin-Gomez, C., Egea-Gamez, R. M., Valera-Calero, J. A., Martin-Casas, P., & Lopez-de-Uralde-Villanueva, I. (2023). Trunk Control Measurement Scale (TCMS): Psychometric Properties of Cross-Cultural Adaptation and Validation of the Spanish Version. *Int J Environ Res Public Health*, *20*(6). https://doi.org/10.3390/ijerph20065144

McHugh, M. L. (2012). Interrater reliability: the kappa statistic. *Biochem Med (Zagreb)*, *22*(3), 276-282.

Mokkink, L. B., Terwee, C. B., Patrick, D. L., Alonso, J., Stratford, P. W., Knol, D. L., Bouter, L. M., & de Vet, H. C. (2010). The COSMIN study reached international consensus on taxonomy, terminology, and definitions of measurement properties for health-related

patient-reported outcomes. *J Clin Epidemiol*, *63*(7), 737-745. https://doi.org/10.1016/j.jclinepi.2010.02.006

Newton, C. R. (2018). Global Burden of Pediatric Neurological Disorders. *Semin Pediatr Neurol*, *27*, 10-15. https://doi.org/10.1016/j.spen.2018.03.002

Novak, I., & Honan, I. (2019). Effectiveness of paediatric occupational therapy for children with disabilities: A systematic review. *Aust Occup Ther J*, *66*(3), 258-273. https://doi.org/10.1111/1440-1630.12573

O'Shea, T. M. (2008). Diagnosis, treatment, and prevention of cerebral palsy. *Clin Obstet Gynecol*, *51*(4), 816-828. https://doi.org/10.1097/GRF.0b013e3181870ba7

Pham, H. P., Eidem, A., Hansen, G., Nyquist, A., Vik, T., & Sæther, R. (2016). Validity and Responsiveness of the Trunk Impairment Scale and Trunk Control Measurement Scale in Young Individuals with Cerebral Palsy. *Phys Occup Ther Pediatr*, *36*(4), 440-452. https://doi.org/10.3109/01942638.2015.1127867

Portney, L. G., & Watkins, M. P. (2009). Foundations of clinical research: Application to practice. *New Jersey: Pearson Education.*

Roberts, P., & Priest, H. (2006). Reliability and validity in research. *Nurs Stand*, *20*(44), 41-45. https://doi.org/10.7748/ns2006.07.20.44.41.c6560

Sæther, R., Helbostad, J. L., Adde, L., Braendvik, S., Lydersen, S., & Vik, T. (2015). The relationship between trunk control in sitting and during gait in children and adolescents with cerebral palsy. *Dev Med Child Neurol*, *57*(4), 344-350. https://doi.org/10.1111/dmcn.12628

Saether, R., Helbostad, J. L., Adde, L., Jorgensen, L., & Vik, T. (2013). Reliability and validity of the Trunk Impairment Scale in children and adolescents with cerebral palsy. *Res Dev Disabil*, *34*(7), 2075-2084. https://doi.org/10.1016/j.ridd.2013.03.029

Silkwood-Sherer, D. J., Killian, C. B., Long, T. M., & Martin, K. S. (2012). Hippotherapy--an intervention to habilitate balance deficits in children with movement disorders: a clinical trial. *Phys Ther*, *92*(5), 707-717. https://doi.org/10.2522/ptj.20110081

Story, D. A., & Tait, A. R. (2019). Survey Research. *Anesthesiology*, *130*(2), 192-202. https://doi.org/10.1097/aln.0000000000002436

Vanbelle, S. (2016). A New Interpretation of the Weighted Kappa Coefficients. *Psychometrika*, *81*(2), 399-410. https://doi.org/10.1007/s11336-014-9439-4

Vitrikas, K., Dalton, H., & Breish, D. (2020). Cerebral Palsy: An Overview. *Am Fam Physician*, *101*(4), 213-220.

APPENDIX

Appendix Table 1

Adjustments HippoTrunC

<table>
<tr><th colspan="5">HippoTrunc</th></tr>
<tr><td colspan="5">Naam, voornaam:
Geslacht:
Geboortedatum:
Testdatum:
GMFCS level:
Video-analyse toestemming: ja/nee
Naam Paard:</td></tr>
<tr><td colspan="2">Hulpmiddelen
~~Per hulpmiddel -1 punt van de totale score~~</td><td colspan="3">~~AFO's: ja - nee~~
Korset: elastisch (-1) - gips (-2)
~~Solo-zit– mate van ondersteuning: geen (0) licht (-1) - matig (-2) - veel (-3)~~
Duo-zit ~~– mate van ondersteuning: licht (-1) – matig (-2) – veel(-3)~~ ja(-1) – nee (0)
Stijgbeugels: ja (-1) - nee (0)</td></tr>
<tr><td colspan="5">Algemeen:
Elke instructie mag 2x uitgelegd worden aan de zorggebruiker.
Voor de test op een bewegend paard wordt afgenomen, worden er twee rondes in de piste toegestaan zodat het kind kan aanpassen aan het bewegend paard.
Met "*recht" wordt in de testbatterij de maximale oprichting bedoeld, die mogelijk is per individu.
De zorggebruikers mogen aangespoord worden recht te zitten en dit zoveel als nodig</td></tr>
<tr><td colspan="5">~~Statische controle~~</td></tr>
<tr><td colspan="5">Item 1. Hoofdcontrole: statisch
Instructie: "Kijk recht voor je terwijl het paard stil staat, gedurende 60 sec"</td></tr>
<tr><td>Score 1</td><td>Score 2</td><td>Score 3</td><td>Score 4</td><td>Score 5</td></tr>
<tr><td>Onmogelijk om het hoofd recht* te houden ~~of heeft nek ondersteuning nodig~~</td><td>Houdt hoofd zelfstandig recht* voor < 10 sec</td><td>Houdt hoofd zelfstandig recht* voor ~~< 60 sec~~
≥ 10 sec en < 30 sec</td><td>Houdt zelfstandig hoofd recht* ~~voor > 60 sec~~
≥ 30 sec en < 60 sec</td><td>Houdt hoofd zelfstandig recht* voor ≥ 60 sec</td></tr>
<tr><td colspan="5">Item 2. Zitbalans: statisch
Instructie: "Zit zo recht mogelijk en kijk voorwaarts, gedurende 60 sec"</td></tr>
<tr><td>Score 1</td><td>Score 2</td><td>Score 3</td><td>Score 4</td><td>Score 5</td></tr>
<tr><td>~~Gebrek aan controle van de romp, heeft veel ondersteuning nodig van een persoon~~
Onmogelijk om romp adequaat op te richten (teveel extensie of flexie), maximale ondersteuning van therapeut nodig.</td><td>~~Houdt romp recht, heeft minimale steun nodig van een persoon~~
Houdt romp zelfstandig recht* >0 en ≤ 10 sec en/of grote voorachterwaartse of zijwaartse bewegingen romp en/of matige ondersteuning van therapeut nodig.</td><td>~~Houdt romp recht zonder steun voor < 30 sec~~
Houdt romp zelfstandig recht* >10 en ≤ 30 sec en/of matige voorachterwaartse of zijwaartse bewegingen romp zichtbaar en/of lichte ondersteuning therapeut nodig.</td><td>~~Houdt romp recht zonder steun voor > 30 sec~~
Houdt romp zelfstandig recht* >30 en ≤ 60 sec, en/of lichte inadequate voorachterwaartse of zijwaartse bewegingen romp zichtbaar.</td><td>Houdt romp zelfstandig recht* voor >60 sec, geen inadequate voorachterwaartse of zijwaartse bewegingen romp zichtbaar.</td></tr>
<tr><td colspan="5">~~Actieve controle~~</td></tr>
<tr><td colspan="5">Item ~~3.~~ 6 Hoofdcontrole: actief</td></tr>
</table>

Instructie: "Kijk recht voor je terwijl het paard beweegt" ~~en "Kijk naar (voorwerp aan de rechterzijde in de ruimte, voorwerp aan de linkerzijde in de ruimte -> 5-10s naar een kant kijken)"~~				
Score 1	**Score 2**	**Score 3**	**Score 4**	**Score 5**
Onmogelijk om het hoofd recht* te houden	Houdt hoofd zelfstandig recht* ~~wanneer het paard beweegt maar niet tijdens rotatie~~ gedurende enkele passen	Houdt hoofd zelfstandig recht* ~~en kan naar één zijde roteren, dit 5 sec aanhouden of naar beide zijden roteren en 5 sec aanhouden zonder dissociatie tussen romp en hoofd~~ gedurende de helft van de lengte	Houdt zelfstandig hoofd recht* ~~en kan hoofd roteren naar beide zijdes en dit 5 sec aanhouden, met dissociatie tussen romp en hoofd~~ gedurende 3/4 e van de volledige lengte	Houdt hoofd zelfstandig recht* gedurende de volledige lengte (16m)
Item ~~4.~~ 5 Zitbalans: actief (normaal tempo) Instructie: "Blijf zo recht mogelijk op het paard zitten terwijl hij voortbeweegt."				
Score 1	**Score 2**	**Score 3**	**Score 4**	**Score 5**
~~Gebrek aan controle van de romp, heeft veel ondersteuning nodig van een persoon~~ Onmogelijk om romp adequaat op te richten (teveel extensie of flexie), maximale ondersteuning therapeut nodig.	~~Houdt romp recht, heeft minimale steun nodig van een persoon~~ Houdt romp zelfstandig recht* gedurende enkele stappen, en/of met grote voorachterwaartse of zijwaartse bewegingen en/of matige ondersteuning van therapeut nodig	~~Houdt romp recht zonder steun voor de helft van de lengte (5m)~~ Houdt romp zelfstandig recht* gedurende de helft van de lengte (5m) en/of met matige voorachterwaartse of zijwaartse bewegingen romp en/of lichte ondersteuning therapeut nodig.	~~Houdt romp recht zonder steun voor de volledige lengte (>5m)~~ Houdt romp zelfstandig recht* gedurende de volledige lengte (10m) met licht inadequate voorachterwaartse of zijwaartse bewegingen romp.	Houdt romp zelfstandig recht* gedurende de volledige lengte (10m), geen inadequate voorachterwaartse of zijwaartse bewegingen romp zichtbaar.
Item ~~5.~~ 7 Zitbalans: actief (versneld tempo) Instructie: "Blijf zo recht mogelijk zitten op het paard terwijl hij sneller gaat stappen."				
Score 1	**Score 2**	**Score 3**	**Score 4**	**Score 5**
~~Gebrek aan controle van de romp, heeft veel ondersteuning nodig van een persoon~~ Onmogelijk om romp adequaat op te richten (teveel extensie of flexie), maximale ondersteuning therapeut nodig.	~~Houdt romp recht, heeft minimale steun nodig van een persoon~~ Houdt romp zelfstandig recht* gedurende enkele stappen, en/of met grote voorachterwaartse of zijwaartse bewegingen en/of matige ondersteuning van therapeut nodig	Houdt romp zelfstandig recht* ~~zonder steun voor~~ gedurende de helft van de lengte (5m) en/of met matige voorachterwaartse of zijwaartse bewegingen van de romp en/of lichte ondersteuning therapeut nodig.	Houdt romp zelfstandig recht* ~~zonder steun voor de volledige lengte (>5m)~~ gedurende de volledige lengte (10m) met licht inadequate voorachterwaartse of zijwaartse bewegingen van de romp.	Houdt romp zelfstandig recht* gedurende de volledige lengte (10m), geen inadequate voorachterwaartse of zijwaartse bewegingen van de romp zichtbaar.
Item ~~6.~~ 3 Zitbalans: actief (reiken)				

Instructie: “Probeer met één hand een wasknijper te grijpen, ga hiervoor zo ver mogelijk richting het oor van het paard en keer daarna terug naar de beginpositie”				
Score 1	**Score 2**	**Score 3**	**Score 4**	**Score 5**
Onmogelijk om voorwaarts te reiken ~~of de arm te strekken richting het oor zonder zijn balans te verliezen of met manuele ondersteuning van de romp~~	Mogelijk om de arm richting het oor te strekken ~~met manuele ondersteuning van de romp,~~ maar geen voorwaartse beweging romp mogelijk zonder het verliezen van de balans, tenzij met maximale ondersteuning therapeut	Mogelijk om de arm te strekken en zelfstandig voorwaarts te reiken (<20cm) ~~richting het oor zonder steun~~ en/of met licht tot matige ondersteuning therapeut (20cm).	~~Mogelijkheid om voorwaarts te reiken (>20cm) op een gecontroleerde manier~~ Mogelijk om zelfstandig arm te strekken en voorwaarts te reiken (20cm), maar moet op het einde van de bewegingsbaan de handen plaatsen om balans te behouden en/of ondersteuning nodig van therapeut om terug te keren tot de beginpositie.	Mogelijk om zelfstandig arm te strekken, voorwaarts te reiken (20cm) en terug te keren naar beginpositie zonder het gebruik van de handen en met behoud van balans.
Item ~~7.~~ 4 Zitbalans: actief (zijwaarts reiken) Instructie: “Probeer met één hand de stok te tikken die naast het paard staat, ga hiervoor zo ver mogelijk richting de stok. Keer daarna terug naar de beginpositie.”				
Score 1	**Score 2**	**Score 3**	**Score 4**	**Score 5**
Onmogelijk om zijwaarts te reiken of de arm te strekken richting de stok ~~zonder zijn balans te verliezen of met manuele ondersteuning van de romp~~	Mogelijk om de arm richting de stok te strekken ~~met manuele ondersteuning van de romp~~ maar geen zijwaartse beweging romp mogelijk zonder het verliezen van de balans, tenzij met maximale ondersteuning therapeut	~~Mogelijk om de arm te strekken (<30cm) richting de stok zonder steun~~ Mogelijk om arm richting de stok te strekken en zelfstandig zijwaarts te reiken (<10cm) en/of met licht tot matige ondersteuning therapeut (10cm)	~~Mogelijkheid om zijwaarts te reiken (>30cm) op een gecontroleerde manier~~ Mogelijk om arm richting de stok te strekken en zelfstandig zijwaarts te reiken (10cm), maar verliest op het einde van de bewegingsbaan de balans en heeft hiervoor ondersteuning nodig van de therapeut om terug te keren tot de beginpositie	Mogelijk om arm richting de stok te strekken, zijwaarts te reiken (10cm) en terug te keren naar beginpositie zonder de hulp van de therapeut.
~~Anticipatoire controle~~				
Item 8. Zitbalans: reactief (slalom) Instructie: “Probeer zo rechtop mogelijk te zitten wanneer het paard beweegt in bochten” (Kinesitherapeut/vrijwilliger beweegt met het paard aan de hand en bepaalt de bewegingen tussen de slalom)				
Score 1	**Score 2**	**Score 3**	**Score 4**	**Score 5**

~~Gebrek aan rompcontrole, ondersteuning van de romp nodig door een persoon~~ Onmogelijk om romp adequaat op te richten (teveel extensie of flexie), maximale ondersteuning therapeut nodig.	~~Houdt romp rechtop maar nood aan rug ondersteuning~~ Houdt romp zelfstandig recht* gedurende enkele stappen, en/of reageert met grote voorachterwaartse of zijwaartse bewegingen van de romp op de verstoringen van het paard en/of matige ondersteuning nodig van de therapeut.	~~Houdt romp rechtop maar kan niet altijd anticiperen op de beweging van het paard/kinesitherapeut~~ Kan zich meerdere keren over de volledige lengte (10m) zelfstandig oprichten, en/of reageert met matige voorachterwaartse of zijwaartse bewegingen van de romp op de verstoringen van het paard en/of lichte ondersteuning nodig van de therapeut.	~~Houdt romp rechtop en anticipeert correct~~ Houdt romp zelfstandig recht* gedurende de volledige lengte (10m) en reageert met licht inadequate voorachterwaartse of zijwaartse bewegingen van de romp op de verstoringen van het paard.	Houdt romp zelfstandig recht* gedurende de volledige lengte (10m), anticipeert correct op de beweging van het paard
Item 9. Zitbalans: reactief (slalom 2) Instructie: "Probeer zo rechtop mogelijk te zitten wanneer het paard beweegt in bochten" (Kinesitherapeut/vrijwilliger beweegt met het paard aan de hand en bepaalt de bewegingen tussen de slalom, deze bochten zijn scherper dan de vorige)				
Score 1	**Score 2**	**Score 3**	**Score 4**	**Score 5**
~~Gebrek aan rompcontrole, ondersteuning van de romp nodig door een persoon~~ Onmogelijk om romp adequaat op te richten (teveel extensie of flexie), maximale ondersteuning therapeut nodig.	~~Houdt romp rechtop maar nood aan rug ondersteuning~~ Houdt romp zelfstandig recht* gedurende enkele stappen, en/of reageert met grote voorachterwaartse of zijwaartse bewegingen van de romp op de verstoringen van het paard en/of matige ondersteuning nodig van de therapeut	~~Houdt romp rechtop maar kan niet altijd anticiperen op de beweging van het paard/kinesitherapeut~~ Kan zich meerdere keren over de volledige lengte (10m) zelfstandig oprichten, en/of reageert met matige voorachterwaartse of zijwaartse bewegingen van de romp op de verstoringen van het paard en/of lichte ondersteuning nodig van de therapeut.	~~Houdt romp rechtop en anticipeert correct~~ Houdt romp zelfstandig recht* gedurende de volledige lengte (10m), met licht inadequate voorachterwaartse of zijwaartse bewegingen van de romp op de verstoringen van het paard.	Houdt romp zelfstandig recht* gedurende de volledige lengte (10m), anticipeert correct op de beweging van het paard
Item 10: Zitbalans: reactief (start-stop) Instructie: "Probeer zo rechtop mogelijk te zitten wanneer het paard ~~beweegt~~ snelheidsveranderingen maakt"				

(Kinesitherapeut/vrijwilliger beweegt met het paard aan de hand en bepaalt de momenten van starten en stoppen)				
Score 1	**Score 2**	**Score 3**	**Score 4**	**Score 5**
~~Gebrek aan rompcontrole, ondersteuning van de romp nodig door een persoon~~ Onmogelijk om de romp adequaat op te richten (teveel extensie of flexie), maximale ondersteuning therapeut nodig.	~~Houdt romp rechtop maar nood aan rug ondersteuning~~ Houdt romp zelfstandig recht* gedurende een deel van de volledige lengte, reageert met grote voorachterwaartse of zijwaartse bewegingen op het starten en stoppen van het paard en/of matige ondersteuning nodig van therapeut.	~~Houdt romp rechtop maar kan niet altijd anticiperen op de beweging van het paard/kinesitherapeut~~ Houdt romp zelfstandig recht* gedurende een deel van de volledige lengte, reageert met matige voorachterwaartse of zijwaartse bewegingen van de romp op het starten en stoppen van het paard en/of lichte ondersteuning nodig van de therapeut.	~~Houdt romp rechtop en anticipeert correct~~ Houdt romp zelfstandig recht* gedurende de volledige lengte (10m), reageert met lichte inadequate voorachterwaartse of zijwaartse bewegingen romp op het starten en stoppen van het paard.	Houdt romp zelfstandig recht* gedurende de volledige lengte (10m), reageert met lichte inadequate voorachterwaartse of zijwaartse bewegingen romp op het starten en stoppen van het paard.
Totale score ~~(/40)~~ (/50)				
Downgrade				
Gecorrigeerde totale score				
Opmerkingen				

Appendix Table 2

HippoTrunC

HippoTrunC

Naam, voornaam: Testdatum:

Geslacht:

Geboortedatum: Naam paard:

GMFCS level: Video-analyse toestemming: ja/neen

Hulpmiddelen		Downgradescore
	Duozit: ja (-1) – nee (0)	
	Stijgbeugels: ja (-1) – nee (0)	
	Korset: elastisch(-1) – gips(-2)	
	TOTAAL	

Algemeen

- Elke instructie mag 2x uitgelegd worden aan de zorggebruiker, vanaf dat moment wordt de zorggebruiker getest.
- Voor de test op een bewegend paard wordt afgenomen, worden er twee rondes in de piste toegestaan zodat het kind kan aanpassen aan het bewegend paard.
- Met "*recht" wordt in de testbatterij de maximale oprichting bedoeld, die mogelijk is per individu.
- De zorggebruikers mogen aangespoord worden recht te zitten en dit zoveel als nodig

Item 1. Hoofdcontrole: statisch

Instructie: "Kijk recht voor je terwijl het paard stil staat, gedurende 60 sec"

Score 1 = Onmogelijk om het hoofd recht* te houden
Score 2 = Houdt hoofd zelfstandig recht* voor <10 sec
Score 3 = Houdt hoofd zelfstandig recht* ≥ 10 sec en <30 sec
Score 4 = Houdt hoofd zelfstandig recht* voor ≥ 30 sec en <60 sec
Score 5 = Houdt hoofd zelfstandig recht* voor ≥ 60 sec

Score item 1	

Item 2. Zitbalans: statisch

Instructie: "Zit zo recht mogelijk en kijk voorwaarts, gedurende 60 sec"

Score 1

Onmogelijk om romp adequaat op te richten (teveel extensie of flexie), maximale ondersteuning van therapeut nodig.

Score 2

Houdt romp zelfstandig recht* >0 en ≤ 10 sec en/of grote voorachterwaartse of zijwaartse bewegingen romp en/of matige ondersteuning van therapeut nodig

Score 3

Houdt romp zelfstandig recht* >10 en ≤ 30 sec en/of matige voorachterwaartse of zijwaartse bewegingen romp zichtbaar en/of lichte ondersteuning therapeut nodig.

Score 4

Houdt romp zelfstandig recht* >30 en ≤ 60 sec, en/of lichte inadequate voorachterwaartse of zijwaartse bewegingen romp zichtbaar

Score 5
Houdt romp zelfstandig recht* voor >60 sec, geen inadequate voorachterwaartse of zijwaartse bewegingen romp zichtbaar.

Score Item 2	

Item 3: Zitbalans: actief (reiken)
Instructie: "Probeer met één hand een wasknijper te grijpen, ga hiervoor zo ver mogelijk richting het oor van het paard en keer daarna terug naar de beginpositie"

Score 1
Onmogelijk om arm te strekken en voorwaarts te reiken.

Score 2
Mogelijk om de arm te strekken, maar geen voorwaartse beweging romp mogelijk zonder het verliezen van de balans, tenzij met maximale ondersteuning therapeut

Score 3
Mogelijk om arm te strekken en zelfstandig voorwaarts te reiken (<20cm) en/of met licht tot matige ondersteuning therapeut (20cm).

Score 4
Mogelijk om zelfstandig arm te strekken en voorwaarts te reiken (20cm), maar moet op het einde van de bewegingsbaan de handen plaatsen om balans te behouden en/of ondersteuning nodig van therapeut om terug te keren tot de beginpositie.

Score 5
Mogelijk om zelfstandig arm te strekken, voorwaarts te reiken (20cm) en terug te keren naar beginpositie zonder het gebruik van de handen en met behoud van balans.

Score item 3	

Item 4: Zitbalans: actief (zijwaarts reiken)
Instructie: "Probeer met één hand de stok te tikken die naast het paard staat, ga hiervoor zo ver mogelijk richting de stok. Keer daarna terug naar de beginpositie"

Score 1
Onmogelijk om arm richting de stok te strekken en zijwaarts te reiken.

Score 2
Mogelijk om de arm richting de stok te strekken, maar geen zijwaartse beweging romp mogelijk zonder het verliezen van de balans, tenzij met maximale ondersteuning therapeut

Score 3
Mogelijk om arm richting de stok te strekken en zelfstandig zijwaarts te reiken (<10cm) en/of met licht tot matige ondersteuning therapeut (10cm)

Score 4
Mogelijk om arm richting de stok te strekken en zelfstandig zijwaarts te reiken (10cm), maar verliest op het einde van de bewegingsbaan de balans en heeft hiervoor ondersteuning nodig van de therapeut om terug te keren tot de beginpositie

Score 5

Mogelijk om arm richting de stok te strekken, zijwaarts te reiken (10cm) en terug te keren naar beginpositie zonder de hulp van de therapeut.	
Score item 4	
Item 5. Zitbalans: actief (normaal tempo) Instructie: “Blijf zo recht mogelijk op het paard zitten terwijl hij voortbeweegt.”	
Score 1 Onmogelijk om romp adequaat op te richten (teveel extensie of flexie), maximale ondersteuning therapeut nodig. Score 2 Houdt romp zelfstandig recht* gedurende enkele stappen, en/of met grote voorachterwaartse of zijwaartse bewegingen en/of matige ondersteuning van therapeut nodig Score 3 Houdt romp zelfstandig recht* gedurende de helft van de lengte (5m) en/of met matige voorachterwaartse of zijwaartse bewegingen romp en/of lichte ondersteuning therapeut nodig. Score 4 Houdt romp zelfstandig recht* gedurende de volledige lengte (10m) met licht inadequate voorachterwaartse of zijwaartse bewegingen romp. Score 5 Houdt romp zelfstandig recht* gedurende de volledige lengte (10m), geen inadequate voorachterwaartse of zijwaartse bewegingen romp zichtbaar.	
Score item 5	
Item 6. Hoofd controle: actief Instructie: “Kijk recht voor je terwijl het paard beweegt”	
Score 1 Onmogelijk om het hoofd recht* te houden Score 2 Houdt het hoofd zelfstandig recht* gedurende enkele passen Score 3 Houdt het hoofd zelfstandig recht* gedurende de helft van de lengte Score 4 Houdt het hoofd zelfstandig recht* gedurende 3/4^{e} van de volledige lengte Score 5 Houdt hoofd zelfstandig recht* gedurende de volledige lengte (16m)	
Score item 6	
Item 7: Zitbalans actief (versneld tempo) Instructie: “Blijf zo recht mogelijk zitten op het paard terwijl hij sneller gaat stappen.”	
Score 1 Onmogelijk om romp adequaat op te richten (teveel extensie of flexie), maximale ondersteuning therapeut nodig. Score 2	

Houdt romp zelfstandig recht* gedurende enkele stappen, en/of met grote voorachterwaartse of zijwaartse bewegingen en/of matige ondersteuning van therapeut nodig

Score 3
Houdt romp zelfstandig recht* gedurende de helft van de lengte (5m), en/of met matige voorachterwaartse of zijwaartse bewegingen van de romp en/of lichte ondersteuning therapeut nodig.

Score 4
Houdt romp zelfstandig recht* gedurende de volledige lengte (10m) met licht inadequate voorachterwaartse of zijwaartse bewegingen van de romp.

Score 5
Houdt romp zelfstandig recht* gedurende de volledige lengte (10m), geen inadequate voorachterwaartse of zijwaartse bewegingen van de romp zichtbaar.

Score item 7	

Item 8. Zitbalans: reactief (slalom)
Instructie: "Probeer zo recht mogelijk te zitten wanneer het paard beweegt in bochten"
(Kinesitherapeut/vrijwilliger beweegt met het paard aan de hand en bepaalt de bewegingen tussen de slalom)

Score 1
Onmogelijk om romp adequaat op te richten (teveel extensie of flexie), maximale ondersteuning therapeut nodig.

Score 2
Houdt romp zelfstandig recht* gedurende enkele stappen, en/of reageert met grote voorachterwaartse of zijwaartse bewegingen van de romp op de verstoringen van het paard en/of matige ondersteuning nodig van de therapeut.

Score 3
Kan zich meerdere keren over de volledige lengte (10m) zelfstandig oprichten, en/of reageert met matige voorachterwaartse of zijwaartse bewegingen van de romp op de verstoringen van het paard en/of lichte ondersteuning nodig van de therapeut.

Score 4
Houdt romp zelfstandig recht* gedurende de volledige lengte (10m) en reageert met licht inadequate voorachterwaartse of zijwaartse bewegingen van de romp op de verstoringen van het paard.

Score 5
Houdt romp zelfstandig recht* gedurende de volledige lengte (10m), anticipeert correct op de beweging van het paard

Score item 8	

Item 9. Zitbalans: reactief (slalom 2)
Instructie: "Probeer zo recht mogelijk te zitten wanneer het paard beweegt in bochten"
(Persoon beweegt met het paard aan de hand en bepaalt de bewegingen tussen de slalom, deze bochten zijn scherper dan de vorige)

Score 1
Onmogelijk om romp adequaat op te richten (teveel extensie of flexie), maximale ondersteuning therapeut nodig.

Score 2
Houdt romp zelfstandig recht* gedurende enkele stappen, en/of reageert met grote voorachterwaartse of zijwaartse bewegingen van de romp op de verstoringen van het paard en/of matige ondersteuning nodig van de therapeut

Score 3
Kan zich meerdere keren over de volledige lengte (10m) zelfstandig oprichten, en/of reageert met matige voorachterwaartse of zijwaartse bewegingen van de romp op de verstoringen van het paard en/of lichte ondersteuning nodig van de therapeut.

Score 4
Houdt romp zelfstandig recht* gedurende de volledige lengte (10m), met licht inadequate voorachterwaartse of zijwaartse bewegingen van de romp op de verstoringen van het paard.

Score 5
Houdt romp zelfstandig recht* gedurende de volledige lengte (10m), anticipeert correct op de beweging van het paard

Score item 9	

Item 10: Zitbalans: reactief (start-stop)
Instructie: "Probeer zo rechtop mogelijk te zitten wanneer het paard snelheidsveranderingen maakt" (Persoon beweegt met het paard aan de hand en bepaalt de momenten van starten en stoppen)

Score 1
Onmogelijk om de romp adequaat op te richten (teveel extensie of flexie), maximale ondersteuning therapeut nodig.

Score 2
Houdt romp zelfstandig recht* gedurende een deel van de volledige lengte, reageert met grote voorachterwaartse of zijwaartse bewegingen op het starten en stoppen van het paard en/of matige ondersteuning nodig van therapeut.

Score 3
Houdt romp zelfstandig recht* gedurende een deel van de volledige lengte, reageert met matige voorachterwaartse of zijwaartse bewegingen van de romp op het starten en stoppen van het paard en/of lichte ondersteuning nodig van de therapeut.

Score 4
Houdt romp zelfstandig recht* gedurende de volledige lengte (10m), reageert met lichte inadequate voorachterwaartse of zijwaartse bewegingen romp op het starten en stoppen van het paard.

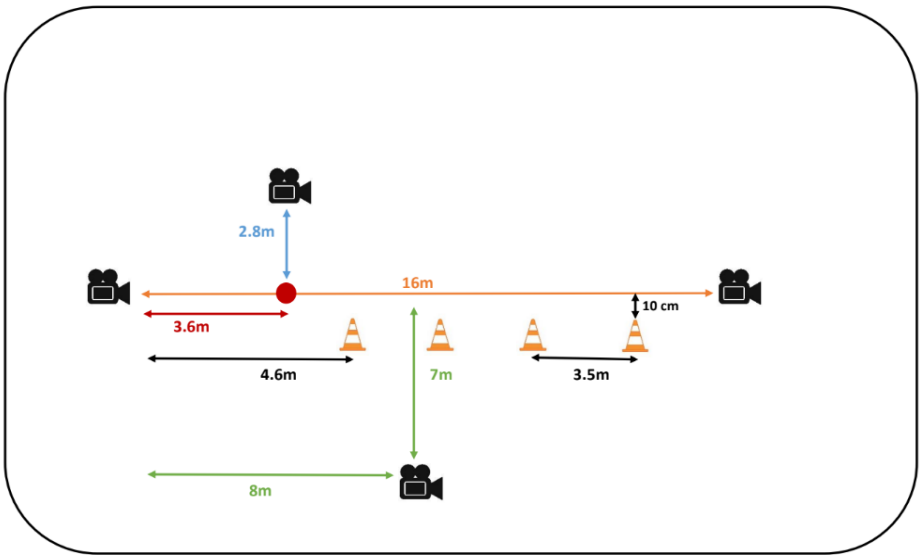
Score 5
Houdt romp zelfstandig recht* gedurende de volledige lengte (10m), anticipeert correct op de beweging van het paard

Score item 10	

Samenvatting HippoTrunC	
Totale score (/50)	
Downgrade	
Gecorrigeerde totale score	
Opmerkingen	

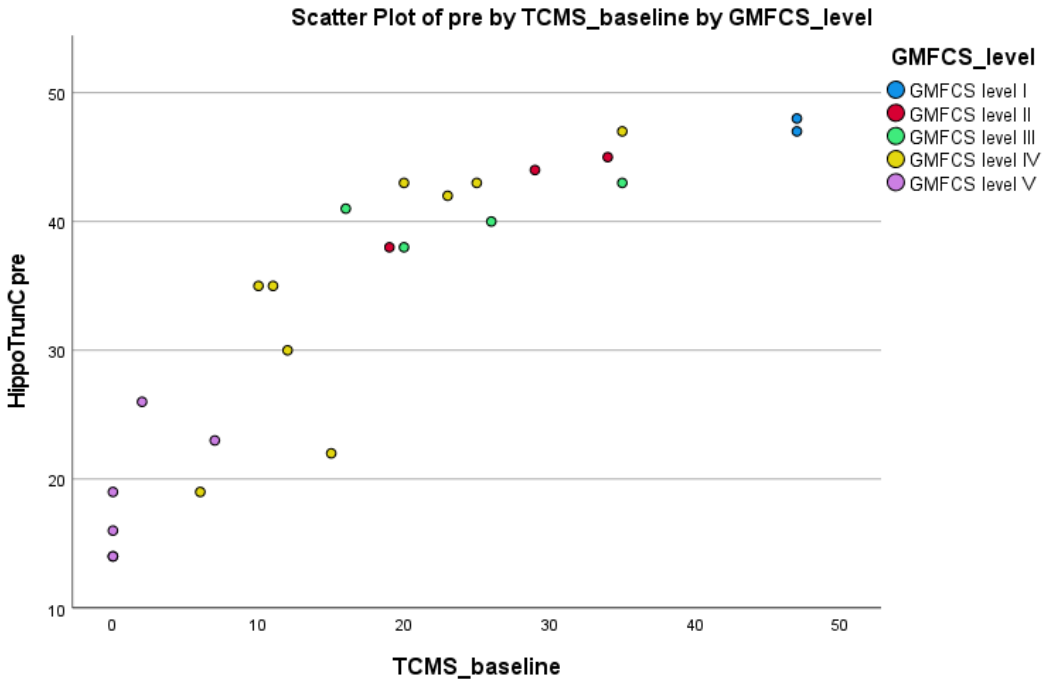
Resultaatniveau	Score	Doel 1	Behaald niveau
Veel minder resultaat dan verwacht	-2	Ik kan mijn romp stabiliseren in de middellijn gedurende 5 sec	x
Iets minder resultaat dan verwacht	-1	Ik kan mijn romp stabiliseren in de middellijn gedurende 15 sec	
Verwacht Resultaat	0	Ik kan mijn romp stabiliseren in de middellijn gedurende 30 sec zonder steun van de therapeut terwijl het paard stilstaat	
Iets meer resultaat dan verwacht	+1	Ik kan mijn romp stabiliseren in de middellijn gedurende 40 sec	
Veel meer resultaat dan verwacht	+2	Ik kan mijn romp stabiliseren in de middellijn gedurende 1 min	

Appendix Figure 1: Example 1 Goal of the Goal Attainment Scale

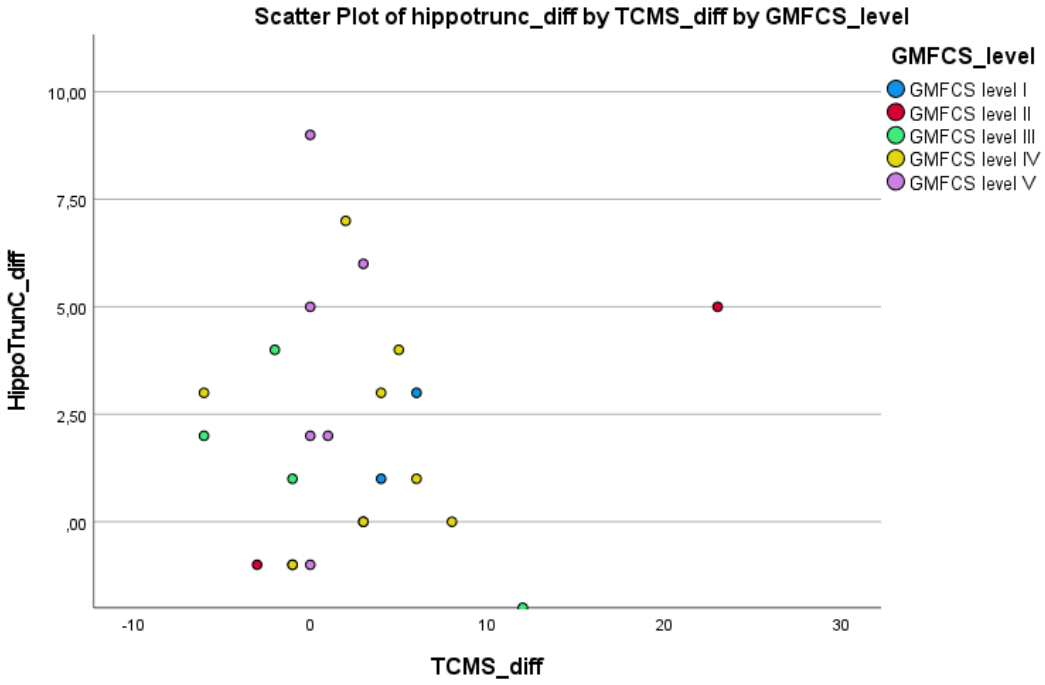


Legend: Red dot = startpoint item 1 till 4; m = meters

Appendix Figure 2: HippoTrunC Course Overview



Appendix Figure 3: Correlation between HippoTrunC Baseline and TCMS Baseline



Appendix Figure 4: Correlation between HippoTrunC Differentiation and TCMS Baseline Differentiation