ISSN 1413-3555
Rev Bras Fisioter, São Carlos, v. 13, n. 3, p. 267-73, May/June 2009


# Validity and intra- and inter-rater reliability of the Observational Gait Scale for children with spastic cerebral palsy

## Validade e confiabilidade intra e interexaminadores da Escala Observacional de Marcha para crianças com paralisia cerebral espástica

Araújo PA[1], Kirkwood RN[2], Figueiredo EM[2]


## Abstract

**Background:** Observational gait assessment is an important clinical approach to the evaluation of gait disorders. Quantitative gait analysis systems provide accurate information, but the high cost of these instruments makes observational analysis more affordable to clinical practice. **Objectives:** To develop an observational gait scale (OGS) for characterizing the gait of children with spastic cerebral palsy (SCP) and to evaluate its validity and reliability criteria in comparison with a computerized motion analysis system representing the gold standard for kinematic gait assessment. **Methods:** Twenty-three videos of children with SCP (9.54 ± 2.22 years) were evaluated by four physical therapists using the OGS, in two sessions. Kinematic data from the ankle/foot complex, knee, hip and pelvis were obtained using the Qualisys Pro-reflex motion analysis system. To establish criterion validity and the intra- and inter-rater reliability, the observational data were compared with motion analysis data, between the two sessions and between the raters. The weighted kappa test was applied to analyze the concordance between the evaluations. **Results:** The OGS presented very good validity for the knee (r=0.64, p<0.05) and good validity for the ankle/foot complex (r=0.59, p<0.05). The intra-rater reliability was very good for the ankle/foot complex (r=0.79, p<0.05), knee (r=0.77, p<0.05) and hip (r=0.73, p<0.05) and good for the pelvis (r=0.59, p<0.05). The inter-rater reliability was very good for the ankle/foot complex (r=0.68, p<0.05) and knee (r=0.65, p<0.05), and good for the hip (r=0.48, p<0.05). **Conclusions:** The OGS demonstrated good reliability and validity for the ankle/foot and knee observations. New observational strategies are needed to improve the psychometric properties of the items relating to the hip and pelvis.


**Key words:** gait; cerebral palsy; scales; reliability; validity.


## Resumo

**Contextualização:** A avaliação observacional da marcha é uma abordagem clínica importante para a avaliação das desordens da marcha. Sistemas de análise quantitativa da marcha oferecem informações acuradas, entretanto o alto custo desses instrumentos tornam a análise observacional mais acessível para a prática clínica. **Objetivos:** Desenvolver uma escala observacional de marcha (EOM) para caracterizar a marcha de crianças com paralisia cerebral espástica (PCE) e testar sua confiabilidade e validade de critério, comparando-a com o sistema computadorizado de análise de movimento, padrão ouro para avaliação cinemática da marcha. **Métodos:** Vinte e três vídeos de crianças com PCE (9,54±2,22 anos) foram avaliados por meio da EOM por quatro fisioterapeutas em duas sessões. Dados cinemáticos do complexo tornozelo/pé, joelho, quadril e pelve foram obtidos usando o sistema de análise de movimento Qualisys Pro-reflex. Para estabelecer a validade de critério e a confiabilidade intra e interexaminadores, os resultados obtidos da EOM foram comparados com os dados do sistema de análise de movimento, entre as duas sessões e entre examinadores. Teste Kappa ponderado foi aplicado para analisar a concordância entre as avaliações. **Resultados:** A EOM apresentou validade muito boa para joelho (r=0,64, p<0,05) e boa para o complexo tornozelo/pé (r=0,59, p<0,05). A confiabilidade intraexaminadores foi muito boa para o complexo tornozelo/pé (r=0,79, p<0,05), joelho (r=0,77, p<0,05) e quadril (r=0,73, p<0,05) e boa para a pelve (r=0,59, p<0,05). A confiabilidade interexaminadores foi muito boa para o complexo tornozelo/pé (r=0,68, p<0,05) e joelho (r=0,65, p<0,05) e boa para o quadril (r=0,48, p<0,05). **Conclusões:** A EOM demonstrou bons índices de confiabilidade e validade para a observação do tornozelo/pé e joelho. Novas estratégias de observação devem ser criadas para melhorar as propriedades psicométricas dos itens relativos ao quadril e pelve.


**Palavras-chave:** marcha; paralisia cerebral; escalas; confiabilidade; validade.

**Received:** 17/09/2008 – **Revised:** 04/12/2008 – **Accepted:** 12/12/2008

[1] Physical Therapist

[2] Department of Physical Therapy, School of Physical Education, Physical Therapy and Occupational Therapy, Universidade Federal de Minas Gerais (UFMG), Belo Horizonte (MG), Brazil

**Correspondence to:** Renata Noce Kirkwood, UFMG – EEFFTO - Departamento de Fisioterapia, Av. Antônio Carlos, 6.627, Pampulha, CEP 31270-901, Belo Horizonte (MG), Brazil, e-mail: renata.kirkwood@gmail.com

# Introduction

Observational analysis is a visual technique used by health professionals to qualitatively describe human gait disorders[1-5]. Its main objective is to infer gait disorders from observational data and document patients' progress during therapeutic intervention[6,7]. There are sophisticated pieces of equipment for quantitative gait analysis[8] that provide accurate information, however the complexity and high cost of these instruments make observational analysis more feasible for clinical practice[9].

The first observational gait scale (OGS), known as Full Body Gait Assessment Form was developed by Rancho Los Amigos Medical Center in California, USA, and published in 1989[1,10]. In 1996, in a study conducted by Greenberg et al.[11], six experienced physical therapists tested the scale watching the gait of twenty-five individuals. The study showed low validity of the scale when compared to the motion analysis system[11]. In 1994, the Physician's Rating Scale (PRS) was created to evaluate knee and foot position during gait in children with spastic cerebral palsy (SCP). There is no record in the literature of the psychometric properties of PRS[13], however several modified versions of this instrument were evaluated for validity and reliability[13-16]. Recently, a new version of PRS, the Visual Gait Assessment Scale (VGAS)[15] was evaluated by two experienced raters using videos from 31 hemiplegic children aged 5 to 17 years. The items with the highest intra-rater reliability were initial contact and foot contact during the stance phase, with greater inter-rater reliability in the parameters of foot contact during stance and heel-off in the terminal stance. The reliability of hip parameters was poor, especially in the swing phase. The authors concluded that the raters' experience influences the reliability results and that more stringent protocols in video collection should be used to improve the validity of the scales[15]. Other studies demonstrate the importance of experience in intra- and inter-rater reliability[17,18].

Several important aspects should be considered to improve the validity and reliability of an observational scale. Video strategies such as freeze frames and slow motion facilitate the observation of gait parameters. Moreover, determining the phase of the gait cycle in which the kinematic parameters must be observed and organizing the scale into categories that indicate only the direction of deviation can favor the reliability and validity of the gait assessment items. Factors such as homogeneous training of observers in the use of the scale and the type of measurement scale (ordinal or nominal) must be clearly standardized to avoid reliability errors and increase measure accuracy. However, these strategies have not been used together in the same observational scale. Therefore, the objectives of this study were to develop an OGS to characterize the gait of children with SCP and test its reliability and criterion validity, comparing it with the computerized motion analysis system, considered the gold standard for kinematic assessment of gait.

# Methods

Data from 18 children with SCP were collected. Data from 5 children were collected bilaterally, resulting in 23 evaluated limbs. Inclusion criteria were clinical diagnosis of SCP, age between 6 and 12 years, ability to walk without human or mechanical aid (Gross Motor Function Classification System – GMFCS – level I or II[19]) and understanding of verbal commands. The presence of motor disorders such as dystonia, ataxia or athetosis was considered criteria for exclusion.

The scale construction followed a script adapted from the model proposed by Benson and Clark[20]. The OGS items were chosen based on the most representative kinematic parameters of gait in children with SCP. Twenty-four items were established for the joints of the ankle/foot (6), knee (5), hip (8) and pelvis (5). The OGS for children with SCP is available at http://www.forusers.com/escala. To access the website, user must request a login and password[21].

For criterion validity and intra- and inter-rater reliability, four physical therapists with experience in the treatment of children with SCP were invited. The criterion validity was determined by agreement between the kinematic gait data, obtained through a motion analysis system, and the OGS data evaluated by the raters using video recordings. For intra-rater reliability, each rater was asked to perform a second gait evaluation of eleven videos selected at random. The second evaluation occurred two weeks after the first to avoid memory bias. Intra-rater reliability was established by agreement between the OGS data completed by the raters in the first evaluation and the data obtained in the second evaluation. Inter-rater reliability was established by agreement between the OGS data of the four raters.

## Procedures

The present study was approved by the Research Ethics Committee of Universidade Federal de Minas Gerais (protocol ETIC 131/05). The guardians of each participant signed a consent form. The kinematic data were obtained using the motion analysis system Qualisys Pro-Reflex - MCU 240 (QUALISYS MEDICAL AB, 411 12 Gothenburg, Sweden) with 5 cameras that recorded the displacement of reflective tracking markers attached to the right and left iliac crest, right and left greater trochanter, lateral and medial femoral epicondyles, clusters attached to the thigh and leg, lateral and medial malleolus and head of the 1st and 5th metatarsals. To capture the images, two digital cameras were positioned perpendicular to (sagittal plane) and in front of (frontal plane) the 8m catwalk where the participants walked. A footswitch was placed on the calcaneus or forefoot region and coupled with a device that lighted when the heel touched the ground. This allowed the synchronization of video and kinematic data.

The four raters were trained on the use of the OGS and on the importance of its items before the evaluations. Each rater received a CD containing the OGS file and the videos of all children in the frontal and sagittal plane. They were free to choose the time and place for OGS application, and the order of video analysis was random.

## Reduction of data

The kinematic data was captured by the acquisition software Qualisys Track Manager 1.9.215. Data from 3 to 5 strides, on average, were analyzed. The data were then transferred to the software Visual3D for the construction of biomechanical models. Footswitch data were used for synchronization and to define the gait cycle from 0 to 100% (from foot contact to the next contact of the same foot). The joint angles for pelvis, hip, knee and ankle were calculated using the Cardan sequence [8]. The graphs of mean angular displacements showed the intervals of the gait phases, which corresponded to the intervals of each OGS item, and showed the maximum and minimum angular peaks. Using normative values available in the literature[22], a confidence interval of 95% was established for each joint angle. This interval was used as normality reference for agreement calculations between the OGS data and kinematic data, thereby establishing the criterion validity.

## Statistical analysis

Weighted kappa (wk) was used to determine the correlation between the kinematic data and OGS data (criterion validity), between the mean OGS data from the four raters in the first and second evaluation (intra-rater reliability) and between the OGS data obtained by the four raters (inter-rater reliability). To interpret the kappa statistics, the following criteria were followed: wk<0.00, no agreement; wk=0.00-0.20, poor agreement; wk=0.21-0.40, fair agreement; wk=0.41-0.60, good agreement; wk=0.61-0.80, very good agreement and wk=0.81-1.00, excellent agreement[22]. The significance level was $\alpha$=0.05.

# Results

Eighteen children participated in the study (11 male and 7 female; 9.54±2.22 years). The body Mass Index (BMI) was 17.18 (3.43) kg/m$^2$. Twelve children had spastic diplegia and 6 had spastic hemiplegia. Among the hemiplegic children, 4 were affected on the left side and 2 on the right side. With regard to GMFCS classification, 11 children were level I and 7 were level II. The children's mean stride velocity was 0.73 m/s (0.25). The four physical therapists had a mean experience of 14.33 (±3.06) years in the treatment of children with SCP. Two work in private clinics and two in rehabilitation hospitals. The mean time for OGS application was 14 minutes, and none of the raters reported difficulties in filling in the scale.

## Validity of test

Items 1, 3 and 4 were related to the ankle/foot complex and were not included in the criterion validity because of the difficulty in calculating the values for this complex with the proposed biomechanical model. Of the 21 items tested, items 7 (wk=0.62) and 9 (wk=0.66) showed very good agreement, followed by items 5 (wk=0.50), 6 (wk=0.55) and 8 (wk=0.52), with good agreement. Items 2 (wk=0.31), 11 (wk=0.27) and 12 (wk=0.23) showed fair agreement. The remaining items had poor to no agreement or were not statistically significant (p>0.05). Regarding intra-rater agreement between OGS and kinematic data, rater 1 showed good agreement (wk=0.45), while the others showed fair agreement (wk between 0.37 and 0.39). The general agreement between OGS data and kinematic data was fair (wk=0.40). The analysis according to joints showed a very good agreement between the OGS and the kinematic knee data (wk=0.64), followed by good agreement in the ankle/foot joint (wk=0.59) and fair (wk=0.23) to poor agreement (wk=0.20) in the pelvis and hip joints, respectively (Table 1).

## Intra-rater reliability

The general agreement index was very good (wk=0.74). Four items showed excellent levels of agreement: item 1 (wk=0.92), item 3 (wk=0.81), item 6 (wk=0.82) and item 16 (wk=0.82). Very good agreement (wk between 0.61 and 0.80) was attained by eleven items, two of them from the ankle joint (items 2 and 4), three from the knee joint (items 7, 8 and 9), four from the hip joint (items 12, 15, 17 and 18) and two from the pelvis joint (items 21 and 23). The eight remaining items showed good to fair agreement (wk=0.60 to 0.26). Rater 1 had an excellent level of agreement (wk=0.91). The other raters had very good agreement (wk=0.74 to 0.62). The intra-rater reliability for OGS joints showed that the ankle/foot (wk=0.79), knee (wk=0.77) and hip (wk=0.73) joints had very good agreement indexes. Only the pelvis joint had a fair agreement index (wk=0.59; Table 2).

**Table 1.** Criterion validity – Agreement between the Observational Gait Scale and the kinematic data per item and joints (N=23).

| Items | rater wk | | | | mean wk | Joint | rater wk | | | | mean wk |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | | | 1 | 2 | 3 | 4 | |
| 2 | 0.26 | 0.24* | 0.41* | 0.39* | 0.31* | | | | | | |
| 5 | 0.51* | 0.45* | 0.43* | 0.70* | 0.50* | Ankle/Foot | 0.49* | 0.51* | 0.66* | 0.69* | 0.59* |
| 6 | 0.47* | 0.61* | 0.58* | 0.57* | 0.55* | | | | | | |
| 7 | 0.62* | 0.74* | 0.37* | 0.70* | 0.62* | | | | | | |
| 8 | 0.48* | 0.36* | 0.61* | 0.64* | 0.52* | | | | | | |
| 9 | 0.71* | 0.72* | 0.67* | 0.49* | 0.66* | Knee | 0.67* | 0.61* | 0.62* | 0.65* | 0.64* |
| 10 | 0.21 | 0.11 | 0.07 | 0.17 | 0.13 | | | | | | |
| 11 | 0.21* | 0.30* | 0.27 | 0.29* | 0.27* | | | | | | |
| 12 | 0.36* | 0.20 | 0.18 | 0.19 | 0.23* | | | | | | |
| 13 | 0.19 | 0.03 | -0.11 | 0.42* | 0.10 | | | | | | |
| 14 | 0.01 | -0.07 | -0.05 | 0.04 | -0.02 | | | | | | |
| 15 | 0.01 | 0.20 | 0.11 | 0.09 | 0.11 | | | | | | |
| 16 | 0.01 | 0.03 | -0.01 | - | 0.00 | Hip | 0.25* | 0.21* | 0.15* | 0.19* | 0.20* |
| 17 | 0.15 | 0.17 | -0.05 | 0.08 | 0.10 | | | | | | |
| 18 | 0.02 | -0.34 | -0.27 | -0.15 | -0.20 | | | | | | |
| 19 | 0.20 | 0.21 | 0.16 | 0.08 | 0.16 | | | | | | |
| 20 | 0.28 | -0.05 | 0.27 | 0.18 | 0.18* | | | | | | |
| 21 | 0.29* | 0.18 | 0.17 | -0.17 | 0.11 | | | | | | |
| 22 | -0.04 | 0.01 | 0.03 | 0.02 | 0.01 | Pelvis | 0.40* | 0.22* | 0.31* | -0.04 | 0.23* |
| 23 | 0.56* | 0.17 | 0.34* | -0.28 | 0.19* | | | | | | |
| 24 | 0.25 | 0.14 | -0.11 | 0.00 | 0.06* | | | | | | |
| | | | | | | Total | 0.45* | 0.39* | 0.37* | 0.39* | 0.40* |

wk=Weighted kappa; *p<0.05.

**Table 2.** Intra-rater reliability – Reliability of the Observational Gait Scale between sessions of the same rater per items and joints (N=23).

| Items | rater wk | | | | mean wk | Joint | rater wk | | | | mean wk |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | | | 1 | 2 | 3 | 4 | |
| 2 | 1.00* | 0.91* | 0.78* | 1.00* | 0.92* | | | | | | |
| 5 | 0.80* | 0.77* | 0.46* | 0.65* | 0.68* | Ankle/Foot | 0.88* | 0.72* | 0.79* | 0.74* | 0.79* |
| 6 | 0.82* | 0.58* | 0.93* | 0.83* | 0.81* | | | | | | |
| 7 | 1.00* | 0.49* | 0.62* | 1.00* | 0.63* | | | | | | |
| 8 | 1.00* | 0.62* | 0.85* | 0.30 | 0.77* | | | | | | |
| 9 | 1.00* | 0.62* | 0.91* | 0.56* | 0.75* | Knee | 0.98* | 0.54* | 0.74* | 0.77* | 0.77* |
| 10 | 1.00* | -0.50 | 0.38 | -0.32 | 0.26* | | | | | | |
| 11 | 0.87* | 0.38 | 0.13 | 0.78* | 0.60* | | | | | | |
| 12 | 0.62* | 0.35 | 0.44 | 0.69* | 0.65* | | | | | | |
| 13 | 0.56* | 0.69* | 1.00* | 0.23 | 0.59* | | | | | | |
| 14 | -0.14 | 0.27 | 0.82* | 0.46 | 0.57* | | | | | | |
| 15 | 1.00* | 0.46* | 0.62* | 0.74* | 0.66* | | | | | | |
| 16 | 1.00* | 0.69* | 1.00* | 0.55* | 0.82* | Hip | 0.84* | 0.64* | 0.79* | 0.60* | 0.73* |
| 17 | 1.00* | 0.00 | 0.56* | 0.66* | 0.67* | | | | | | |
| 18 | 0.42* | 0.65* | 0.66* | 0.00 | 0.53* | | | | | | |
| 19 | 1.00* | 0.80* | 0.31 | 0.60* | 0.62* | | | | | | |
| 20 | 1.00* | 0.00 | 0.64* | -0.11 | 0.36* | | | | | | |
| 21 | 1.00* | 0.41 | 0.49* | 0.50* | 0.63* | | | | | | |
| 22 | 1.00* | -0.27 | 0.50* | 0.67* | 0.43* | Pelvis | 0.97* | 0.31* | 0.61* | 0.43* | 0.59* |
| 23 | 0.91* | 0.34 | 0.59* | 0.92* | 0.70* | | | | | | |
| 24 | 1.00* | 0.00 | 0.34 | -0.49 | 0.34* | | | | | | |
| | | | | | | Total | 0.91* | 0.62* | 0.74* | 0.68* | 0.74* |

wk=Weighted kappa; *p<0.05.

## Inter-rater reliability

Six items had very good agreement: item 1 (wk=0.73), item 2 (wk=0.66), item 4 (wk=0.66), item 5 (wk=0.61), item 6 (wk=0.65) and item 9 (wk=0.68). One item from the ankle/foot complex (item 3), two from the knee (items 7 and 8) and three from the hip joint (items 15, 18 and 19) had good agreement (wk between 0.56 and 0.41). In the twelve remaining items, the agreement ranged from fair to poor (wk=0.56 to 0.14) or was not statistically significant (p>0.05). Regarding the results by joints, the ankle/foot complex and knee had very good agreement indexes (wk=0.68 and 0.65 respectively). The hip joint had good agreement (wk=0.48), followed by the pelvis with fair agreement (wk=0.30; Table 3).

# Discussion

The purpose of the present study was to develop an OGS with a good reliability and validity index for children with SCP. The OGS was shown to be simple and easy to apply because the average time spent filling it in was short and the raters did not report difficulties in identifying the items. The literature indicates that the use of videos in the sagittal and frontal planes, rater training and item organization into categories that indicate only the direction of deviation appear to be effective strategies to improve validity and reliability of OGSs[13,23]. The results of the present study confirm these hypotheses because, after the implementation of these strategies, some of the investigated items showed high agreement indexes when compared to previous studies[11,13,15,16].

Items that obtained better criterion validity were knee flexion/extension during the heel-strike phase (item 7) and during the mid-stance and pre-swing phases (item 9). Previous studies indicate a fair agreement for the knee and ankle/foot joints[13,15,16], and the results of the present study support these findings. This result may be associated with the greater range of motion (ROM) of these joints in the sagittal plane and in the stance phase. The hip and pelvis joints did not have a fair agreement, as observed in other studies[15,17,24]. This difficulty can be attributed to the synchronized movement of the pelvis and hip and to the limited ROM of these joints, especially in the frontal and transverse planes. Therefore, new strategies must be developed to facilitate the visualization of these segments in the gait cycle.

Regarding the planes of movement, the present study points to the best agreement of items in the sagittal plane. Of the eight items that obtained better results, seven are from sagittal plane. This was expected because ROM is greater in this plane, allowing easier identification of changes. Moreover, the deviations of the sagittal plane describe the most common gait patterns in children with SCP, such as crouch and equine gait. This probably makes raters more familiar with deviations in that plane, particularly in the ankle/foot and knee complex. Nevertheless, two items (2 and 11) did not obtain a fair agreement. Item 11 evaluates knee flexion during the swing phase. It is more difficult to visually follow the trajectory of the lower limb during the swing phase because the ground cannot be used as a reference point. In fact, of seven items evaluated during the swing phase, only item 6 obtained good results. In the frontal plane, no result was statistically significant and, in the transverse plane only one item (6) obtained good agreement. Other studies showed similar results, indicating that the observational analysis in these planes is very limited in children with SCP due to the frequent combination of rotational changes with changes in lower limb adduction[25-28].

Overall, the intra-rater reliability was considered very good. Two raters were especially consistent. Rater 1 had excellent agreement, and rater 3, very good agreement. The two other raters also showed good agreement in the observations. This suggests that the OGS items are sufficiently clear, particularly in the ankle/foot and knee joint, where the results were very good. Among the knee items, only item 10 did not achieve a fair agreement, confirming previous findings that

**Table 3.** Inter-rater reliability – Reliability of the Observational Gait Scale between raters per items and joints (N=23).

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| Ankle/Foot | Items | 1 | 2 | 3 | 4 | 5 | 6 | | |
| | mean wk | 0.73* | 0.66* | 0.48* | 0.66* | 0.61* | 0.65* | | |
| Knee | Items | 7 | 8 | 9 | 10 | 11 | | | |
| | mean wk | 0.51* | 0.51* | 0.68* | 0.34* | 0.32* | | | |
| Hip | Items | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 |
| | mean wk | 0.19* | 0.13 | 0.39* | 0.55* | 0.12 | 0.31* | 0.41* | 0.56* |
| Pelvis | Items | 20 | 21 | 22 | 23 | 24 | | | |
| | mean wk | 0.1 | 0.27* | 0.28* | 0.27* | 0.14* | | | |
| Joints | | Ankle/Foot | Knee | Hip | Pelvis | Total | | | |
| | mean wk | 0.68* | 0.65* | 0.48* | 0.30* | 0.56* | | | |

wk=Weighted kappa; *p<0.05.

the frontal plane is difficult to be evaluated by observation. The hip joint obtained very good agreement, but there was more variability in agreement for each item, especially those in the frontal plane. The greatest variability between raters occurred in the pelvis joint, which corroborates previous studies[24, 25]. De Luca et al.[25] highlighted the difficulty in evaluating pelvic movements, especially in the frontal plane, because of the relationship between the position of the trunk and of the lower extremity during gait[25]. The items on the sagittal plane obtained very good intra-rater agreement, confirming that visualization is easier on this plane, including the pelvis. The items evaluated in the present study had a higher intra-rater agreement index than previous studies[13,15,16], indicating that the strategies adopted in OGS development were effective. Other previously studied scales (OGS and VGAS)[13,17,18] proposed items categorized by angular intervals and had more categories than the present scale. This factor may have influenced the results of the present study, increasing the intra-rater reliability of the OGS.

The inter-rater reliability had a good agreement index. The inter-rater agreement of the ankle/foot and knee joints was very good. Nevertheless, the knee items had lower agreement. Only item 9 had agreement greater than 0.60, in contrast with intra-rater results, in which all knee items had an agreement greater than or equal to 0.60, except for item 10, which is measured in frontal plane. This suggests that, although consistent, raters evaluate these items differently. This hypothesis reinforces the need for more extensive training in the use of OGS to achieve homogeneous observations. The results for ankle/foot and hip joint are similar. The pelvis, as expected, did not have good inter-rater agreement.

To improve the consistency of pelvis and hip observations, we would recommend horizontal and vertical axes of reference in the environment, such as a symmetric graph to facilitate joint position estimation. The raters' experience seems to have contributed to the high consistency of intra-rater data. Although prior training was conducted, the inter-rater agreement was lower than the intra-rater, suggesting that more extensive training is needed, especially for less experienced observers.

The results of the present study contribute to the development of observational gait analysis in children with SCP. The strategies used to develop the OGS, i.e. items with only three categories indicating the direction of deviation, simultaneous video recording and motion analysis, freeze frame and slow motion as well as observer training, seem to have effectively contributed to the better validity and reliability of the OGS when compared to other observational scales. The OGS has been shown to be simple and easy to apply. It is worth noting that the rater's clinical experience contributes to a more reliable result, and that the scale's validity and intra- and inter-rater reliability was more satisfactory for the observations in the sagittal plane.

## Acknowledgments

The authors wish to thank the participants and their family members, the Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) and the Fundação de Amparo à Pesquisa do Estado de Minas Gerais (FAPEMIG) for their financial support, and ForUsers Technology Ltda for the development of the scale in electronic format.

## References

1. Toro B, Nester C, Farren P. A review of observational gait assessment in clinical practice. Physiother Theory Pract. 2003;19(3):137-49.

2. Eastlack ME, Arvidson J, Snyder-Mackler L, Danoff JV, McGarvey CL. Interrater reliability of videotaped observational gait-analysis assessments. Phys Ther. 1991;71(6):465-72.

3. Krebs DE, Edelstein JE, Fishman S. Reliability of observational kinematic gait analysis. Phys Ther. 1985;65(7):1027-33.

4. Saleh M, Murdoch G. In defence of gait analysis. Observation and measurement in gait assessment. J Bone Joint Surg Br. 1985;67(2):237-41.

5. Malouin F. Observational gait analysis. In: Craik R, Oatis CA (editors). Gait analysis: theory and application. St Louis: Mosby; 1995. p. 112-29.

6. Gage JR. The treatment of gait problems in cerebral palsy. 2ª ed. London: Mac Keith; 2004.

7. Perry J. Gait analysis: normal and pathological function. Thorofare: Slack Incorporated; 1992.

8. Cappozzo A, La Croce U, Leardini A, Chiari L. Human movement analysis using stereophotogrammetry. Part 1: theoretical background. Gait Posture. 2005;21(2):186-96.

9. Toro B, Nester CJ, Farren PC. The status of gait assessment among physiotherapists in the United Kingdom. Arch Phys Med Rehabil. 2003;84(12):1878-84.

10. Rancho Los Amigos National Rehabilitation Center. Observational gait analysis handbook. 4a ed. Downey: Los Amigos Research and Education Institute Incorporated; 2001.

11. Greenberg M, Gronley J, Perry J, Lawthwaite R. Concurrent validity of observational gait analysis using the vicon motion analysis system. Gait Posture. 1996;4(2):167-8.

12. Koman LA, Mooney JF 3rd, Smith BP, Goodman A, Mulvaney T. Management of spasticity in cerebral palsy with botulinum-A toxin: report of a preliminary randomized double-blind trial. J Pediatr Orthop. 1994;14(3):299-303.

13. Mackey AH, Lobb GL, Walt SE, Stott NS. Reliability and validity of the Observational Gait Scale in children with spastic diplegia. Dev Med Child Neurol. 2003;45(1):4-11.

14. Corry IS, Cosgrove AP, Duffy CM, Mcneill S, Taylor TC, Graham HK. Botulinum toxin A compared with stretching casts in the treatment of spastic equinus: a randomised prospective trial. J Pediatr Orthop. 1998;18(3):304-11.

15. Dickens WE, Smith MF. Validation of a visual gait assessment scale for children with hemiplegic cerebral palsy. Gait Posture. 2006;23(1): 78-82.

16. Wren TA, Rethlefsen SA, Healy BS, Do KP, Dennis SW, Kay RM. Reliability and validity of visual assessments of gait using a modified physician rating scale for crouch and foot contact. J Pediatr Orthop. 2005;25(5):646-50.

17. Brown CR, Hillman SJ, Richardson AM, Herman JL, Robb JE. Reliability and validity of the Visual Gait Assessment Scale for children with hemiplegic cerebral palsy when used by experienced and inexperienced observers. Gait Posture. 2008;27(4):648-52.

18. Ong AM, Hillman SJ, Robb JE. Reliability and validity of the Edinburgh Visual Gait Score for cerebral palsy when used by inexperienced observers. Gait Posture. 2008;28(2):323-6.

19. Rosenbaum PL, Walter SD, Hanna SE, Palisano RJ, Russell DJ, Raina P, et al. Prognosis for gross motor function in cerebral palsy: creation of motor development curves. JAMA. 2002;288(11):1357-63.

20. Benson J, Clark F. A guide for instrument development and validation. Am J Occup Ther. 1982;36(12):789-800.

21. forusers.com [página na internet], Brasil. ForUsers Tecnologia Ltda, 2002.; [revisão em Janeiro, 2009]. Disponível em: http://www.forusers.com/escala/.

22. Kirtley C. CGA Normative Gait Database [homepage na internet]. Hong Kong; Polytechnic University. [Atualizada em: 2007; acesso em: 12/11/2005]. Disponível em: http://www.univie.ac.at/cga/data/index.html.

23. Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics. 1977;33(1):159-74.

24. Kawamura CM, de Morais Filho MC, Barreto MM, de Paula Asa SK, Juliano Y, Novo NF. Comparison between visual and three-dimensional gait analysis in patients with spastic diplegic cerebral palsy. Gait Posture. 2007;25(1):18-24.

25. de Luca PA, Davis RB 3rd, Ounpuu S, Rose S, Sirkin R. Alterations in surgical decision making in patients with cerebral palsy based on three-dimensional gait analysis. J Pediatr Orthop. 1997;17(5):608-14.

26. Viehweger E, Helix M, Jacquemier M, Scavarda D, Rohon MA, Scorsone-Pagny S. Application of the Edinburgh visual gait score: interobserver and intraobserver reliability. J Bone Joint Surg. 2005;87B Suppl 1:S69.

27. Read HS, Hazlewood ME, Hillman SJ, Prescott RJ, Robb JE. Edinburgh visual gait score for use in cerebral palsy. J Pediatr Orthop. 2003;23(3):296-301.

28. Hillman SJ, Hazlewood ME, Loudon IR, Robb JE. Can transverse plane rotations be estimated from video tape gait analysis? Gait Posture. 1998;8(2):87-90.