

Article

Test-Retest, Inter-Rater and Intra-Rater Reliability for Spatiotemporal Gait Parameters Using SANE (an eaSy gAit aNalysis systEm) as Measuring Instrument

Betsy D. M. Chaparro-Rico  and Daniele Cafolla * 

Biomechatronics Lab, IRCCS Neuromed, 86077 Pozzilli (IS), Italy; betsychaparro@hotmail.com

* Correspondence: contact@danielecafolla.eu

Received: 29 June 2020; Accepted: 17 August 2020; Published: 20 August 2020



Abstract: Studies have demonstrated the validity of Kinect-based systems to measure spatiotemporal parameters of gait. However, few studies have addressed test-retest, inter-rater and intra-rater reliability for spatiotemporal gait parameters. This study aims to assess test-retest, inter-rater and intra-rater reliability of SANE (eaSy gAit aNalysis system) as a measuring instrument for spatiotemporal gait parameters. SANE comprises a depth sensor and a software that automatically estimates spatiotemporal gait parameters using distances between ankles without the need to manually indicate where each gait cycle begins and ends. Gait analysis was conducted by 2 evaluators for 12 healthy subjects during 4 sessions. The reliability was evaluated using Intraclass Correlation Coefficients (ICC). In addition, the Standard Error of the Measurement (SEM), and Smallest Detectable Change (SDC) was calculated. SANE showed from an acceptable to an excellent test-retest, inter-rater and intra-rater reliability; test-retest reliability ranged from 0.62 to 0.81, inter-rater reliability ranged from 0.70 to 0.95 and intra-rater ranged from 0.74 to 0.92. The subject behavior had a greater effect on the reliability of SANE than the evaluator performance. The reliability values of SANE were comparable with other similar studies. SANE, as a feasible and markerless system, has large potential for assessing spatiotemporal gait parameters.

Keywords: gait analysis; spatiotemporal parameters; depth sensor; Kinect; neurorehabilitation; reliability; validity; test-retest reliability; inter-rater reliability; intra-rater reliability

1. Introduction

The restoration of autonomous and functional ambulation is a major priority. A gait study can identify variations and movements impairments that can help to make therapeutic decisions and to estimate the recovery status. Clinical gait analysis may also help to distinguish between disease entities and to determine the risk of disease or injury [1–4]. Base clinical gait evaluations are primarily observational or gait speed-based and they are ideal for measuring, tracking and evaluating a wide range of population functions and general health [5]. However, base clinical gait evaluations do not have the precision or the richness of the data of instrumented methods which are necessary for comprehensive gait analysis on the kinematic and spatiotemporal aspects of the gait cycle [3,4]. However, instrumental gait analysis calls for expensive instruments, which in clinical settings are not always available.

Precise, non-intrusive and cost-efficient clinical gait analysis devices have many diagnosis, tracking, treatment and recovery uses [3,6]. Such uses include early detection and evaluation, assessing drug efficacy in the home or even improving care directly. A number of methods for gait analysis have been suggested in the state of art. Marker systems typically use infrared (IR) cameras and markers to be placed on the subject's body. These systems are reliable, but also very expensive and

unsustainable. In addition, before each capture session, passive or active markers have to be placed on the body correctly.

The use of wearable sensors has been proposed in recent studies [7]. These systems are suitable for ambulatory measurements at home since they are small, easy, mobile and cheaper. Wearable sensors need to be positioned correctly and safely. In addition, the weight, noise and signal drift must be taken into account. Each sensor is normally limited to measuring very few gait properties and thus, a number of sensors are needed to obtain a complete analysis. In addition, these sensors entail adjustments to the subject's daily routine. Furthermore, they often need maintenance in the form of batteries, data upload and sanitation. To avoid these inadequacies, in the context of gait analysis, individuals can be identified using single or multiple video cameras [8,9].

Three-dimensional (3D) depth sensing may provide useful healthcare data such as location, posture and movement of patients [10]. In addition, it can collect 3D body physics measurements [10]. This ability to produce quantitative data helps meet the clinician's need to make decisions based on accurate measures and promote cost-effective custom medical practice [11]. A well-known 3D depth sensor is Microsoft Kinect (Microsoft Co., Redmond, WA, USA), named Kinect v2 in its latest version. Kinect is an input system designed for Xbox video game console computer gaming. The sensor allows a user to communicate in virtual reality through body motion, hand movements and speech commands. The sensor uses color camera, infrared (IR) emitter and IR sensor to create a three-dimensional (3D) image containing a cloud with a location and surface of more than 200,000 points representing the object as coordinates x , y and z . Kinect can be used to promote a healthy life by tracking patterns of activity, by collecting health data, or by detecting early warning signs in elderly or poor health people [12]. In addition, Kinect applications are used for disease control and monitoring as well as to determine patient attitude and motion [10,11,13–15], and to design rehabilitation exercises and therapy devices [16–19].

In previous works, we applied a supervised approach to the learning process to determine gait parameters automatically and precisely using a simulated 3D skeleton [20]. It helped to go beyond normal stage parameters. The proposed method was compared with one commercial vision system used in protocols for clinical trials. The findings showed that the method is equivalent to the commercial system. Currently, the gait analysis system described in [20] is an end-product ready for end-users (noncommercial yet) that has been named SANE (eaSy gAit aNalysis systEm). SANE system is a low-cost, non-intrusive solution which can accurately measure a wide range of gait parameters via a depth sensor and a connected software working together. One of the main advantages of SANE is that the parameter calculation is carried out automatically, that is, it is not necessary to manually indicate where each gait cycle begins and ends since SANE detects it automatically using the distance between ankles. SANE uses a Microsoft Kinect v2 sensor (Microsoft Co., Redmond, WA, USA) as the depth sensor.

Inter-method reliability assesses the agreement between measures of different methods or instruments. Therefore, a new method or instrument can be compared with existing methods calculating the inter-method reliability [21]. Recent studies have addressed the inter-method reliability of Kinect-based systems for gait analysis and promising results have been found for spatiotemporal parameters with respect to other existing methods/systems, as demonstrated in the detailed review about validity of the Kinect for gait assessment in [1]. For example, authors in [22] found a good to excellent agreement between the measures of spatiotemporal gait parameters (gait speed, cadence, step length, stride length, step width, step time and stride time) of a Kinect V2-based system, compared with a 60 Hz Optotrak System 3D motion analysis (Northern Digital Inc., Waterloo, ON, Canada) and a standardized clinical method namely 10 m walking test (10 MWT). As well, the authors in [10] found excellent agreement between the spatiotemporal measures namely gait speed, step length and stride length of a Kinect-based system compared with a 120 Hz VICON 3D motion analysis (VICON, UK). Other comparable results are described in detail in the focused review about the validity of Kinect for gait assessment in [1], when the inter-method reliability is addressed for Kinect-based systems,

generally at the same time the intra-rater reliability is assessed [10,23]. However, few studies include test-retest reliability, the intra-rater reliability and inter-rater reliability of Kinect-based systems when it is used for assessing spatiotemporal parameters of the human gait [24–26]; particularly, inter-rater reliability is almost never addressed. However, the three—test-retest, inter-rater and intra-rater reliability—are important to assess the agreement between measures of an instrument. The test-retest reliability allows to assess the agreement between measures obtained by one evaluator that tests a same group of subjects at different times (when giving the same task to the same subjects two or more times). The inter-rater reliability allows to assess the agreement between the measures obtained by two different evaluators that test the same group of subjects (when giving the same observations to two or more evaluator). Intra-rater reliability allows to assess the agreement between repeated measures obtained by one evaluator that tests a same group of subjects (when giving repeatedly the same observations to one evaluator) [27–32]. The training of the evaluator and the standardization of the task influence the inter-rater and intra-rater reliability. On the other hand, a highly dependent on the situation or on the subject's condition influences the test-retest reliability [27]. This study aims to assess the test-retest, inter-rater and intra-rater reliability of SANE for spatiotemporal parameter of the gait. The estimation of test-retest, inter-rater and intra-rater reliability of SANE were carried out in order to verify through the agreement between measures that SANE is a feasible markerless system for assessing spatiotemporal gait parameters.

2. Materials and Methods

2.1. Participants

This study was carried out at the Biomechanics Lab of the IRCCS Neuromed-Mediterranean Neurological Institute with the participation of 12 healthy subjects: 7 males and 5 females. The subjects voluntarily participated in this study. The subjects do not have any physical contact with any instrument or person during this study since it is used a non-invasive and non-contact method. In addition, it is important to mention that during this study the subjects carried out an activity that is part of their daily live routine: walking a few steps. The inclusion criteria were: absence of any neurological problems and absence of diseases or injuries that can affect the gait cycle. Inclusion or exclusion criteria were not imposed regarding age, height, weight or gender. Characteristics of the subjects included in the study are shown in Table 1. The age of the subjects ranges from 28 to 55 years, the weight from 64 to 100 kg, and the height from 1.58 to 1.78 m.

Table 1. Characteristics of the subjects included in the study, $n = 12$.

Subject	Age (Years)	Weight (kg)	Height (m)	Sex
1	47	70	1.77	M
2	31	80	1.77	M
3	29	58	1.70	M
4	35	79	1.75	M
5	32	85	1.65	M
6	31	64	1.70	F
7	55	82	1.58	F
8	31	70	1.78	F
9	28	65	1.70	M
10	45	100	1.80	M
11	29	92	1.82	M
12	42	82	1.59	F

2.2. Description of SANE

SANE (eaSy gAit aNalysis systEm) [20], is a low-cost and smart system that non-intrusively can accurately measure a wide range of gait parameters. SANE is composed by a depth sensor (Kinect V2)

and a user interface with an algorithm that automatically calculates gait parameters using the tracked points of the human body. That is, it is not necessary to manually indicate where each gait cycle begins and ends since SANE automatically detects each gait cycle. SANE, through the use of the depth sensor, is able to record the position in real time of various key points of the human body that can reconstruct a virtual skeleton. Using the key points, SANE automatically calculate the gait parameters and plots the trend of points of the hip, knee, ankle and pelvis, as well as the foot distance and ankle distance. SANE was developed at Biomechanics Lab of the IRRCS Neuromed-Mediterranean Neurological Institute. SANE arises from the need to have a non-contact and low-cost system for gait analysis so that medical specialists can evaluate the gait in patients that, due to their health condition, it is impossible to use markers or invasive devices on them. A system like SANE (that does not require calibration, is low-cost, uses only one sensor, does not require a structured environment, can be used at home, does not require any physical contact between none of the parts of the system and the subject, has a dedicated graphic interface, does not require markers or references points and automatically calculates the spatiotemporal gait parameters without the need to manually indicate where each gait cycle begins and ends) is not commercially available for spatiotemporal gait analysis.

A Flow-chart with the acquisition procedure using SANE is shown in Figure 1. First, the subject is asked to start walking. The specialist starts the acquisition and the depth sensor (Kinect V2) start to communicate with the software for data collection, meanwhile the trend of point coordinates of the hip, knee, ankle and pelvis, as well as the foot distance and ankle distance are plotted. The real-time plots are useful for the specialist to check both if the dataflow is going and if there is some problem in the subject walking. When the subject arrives to the end of the prescribed path the acquisition is stoped and the data is saved to the subject folder inside the section folder. When the walking section is finished and the data has been saved, it is possible to compute all the mean data and the gait parameters in terms of gait cycle duration, step duration, cadence, gait cycle length and mean velocity. Multiple trials can be saved. However, one trial is enough to calculated the gait parameters.

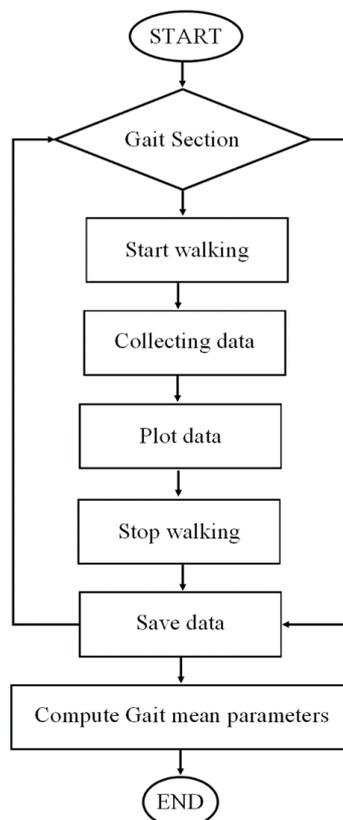


Figure 1. Gait analysis procedure flow-chart.

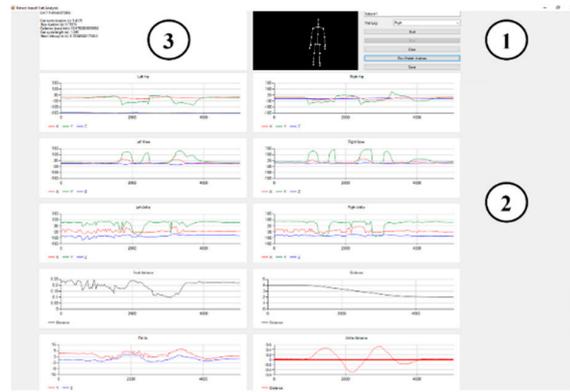
SANE is able to automatically detect a gait cycle using ankle distances. If the first step is done using the left foot the gait cycle is between two subsequent local minima of the function, while if the first step is done with the right foot, the gait cycle is between two subsequent local maxima of the ankle distance plot.

The specialist is helped and guided through a user-friendly interface programmed in C#, Figure 2. The Area marked with 1 of the user interface presents the virtual skeleton view, a text field to input the subject name and a drop-down menu to select starting foot, important for the automatic step detection. The buttons in Area 1 are used to start or stop motion recording clear real-time plots or save the walking trial. Every parameter of every single walk is saved. When the user clicks on the button named "Run Matlab analysis", an instance of Matlab 2017b (MathWorks Inc., Natick, MA, USA) is called for, calculating the gait parameters. The area marked with 2 is where the collected data can be seen plotted in real-time. Having detected the gait cycle, the system operates the cutting smoothing and alignment operations accordingly. At the end of all the mentioned operations, the mean gait parameters are shown on the interface in the Area 3 in terms of gait cycle duration, step duration, cadence, gait cycle length and mean velocity.

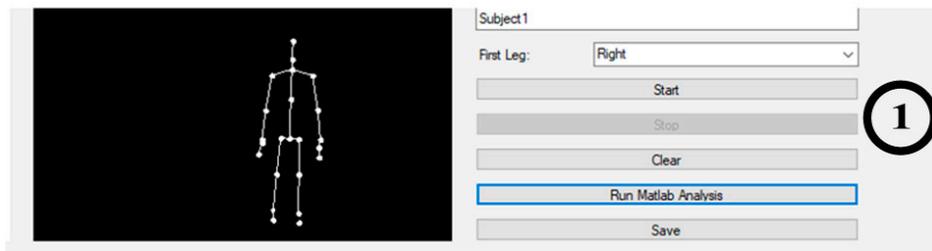
2.3. Procedure

Gait parameters were assessed by SANE for the 12 subjects: cycle duration, step duration, cadence, cycle length and mean velocity. Figure 3 shows the experimental layout used to carry out the data collection. A walking path of 2.3 m was pointed with tape on the floor. The starting and ending of the walking path were pointed with perpendicular lines. In order to obtain a frontal view of the subjects, the depth sensor of SANE was placed in front of the walking path at a height of 0.9 m and a distance of 1.7 m from the ending point. A trial consisted of the subject having to walk from the starting to the ending of the path in Figure 3, so that the subject was positioned in front of the depth sensor and walked towards the depth sensor. While the subject was walking, SANE collected the required data to calculate the gait parameters as explained in Section 3.2. Each subject walked with his/her natural speeds and gait patterns.

The test protocol consisted of four sessions. During each session, each subject performed 10 trials that SANE used to calculate the mean of the gait parameters. So that, mean of gait parameters was obtained for each subject during each session. All subjects participated in all sessions. Session 1, Session 2 and Session 3 were carried out successively in the same day and Session 4 after a seven-day interval. Two trained evaluators (Eval 1 and Eval 2) were part of the experiment. On the first day, Session 1 and Session 2 were carried out successively by the first evaluator (Eval 1) and Session 3 was carried out by the second evaluator (Eval 2), Table 2. Seven days later, Session 4 was carried out by the first evaluator (Eval 1), Table 2. Both evaluators had at least 8 years of experience in motion analysis and they are trained in the use of SANE. Each evaluator explained the procedure to each subject before each session. During each session, each evaluator indicated to the subjects when they should start walking and when they could return to their starting position.



(a)

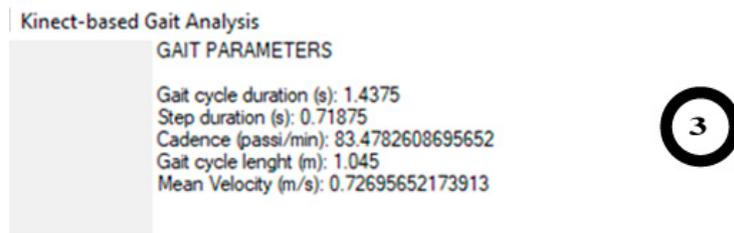


(b)



(c)

Figure 2. Cont.



(d)

Figure 2. SANE (eaSy gAit aNalysis system) interface: (a) overview; (b) zoom-Area 1; (c) zoom-Area 2; (d) zoom-Area 3.

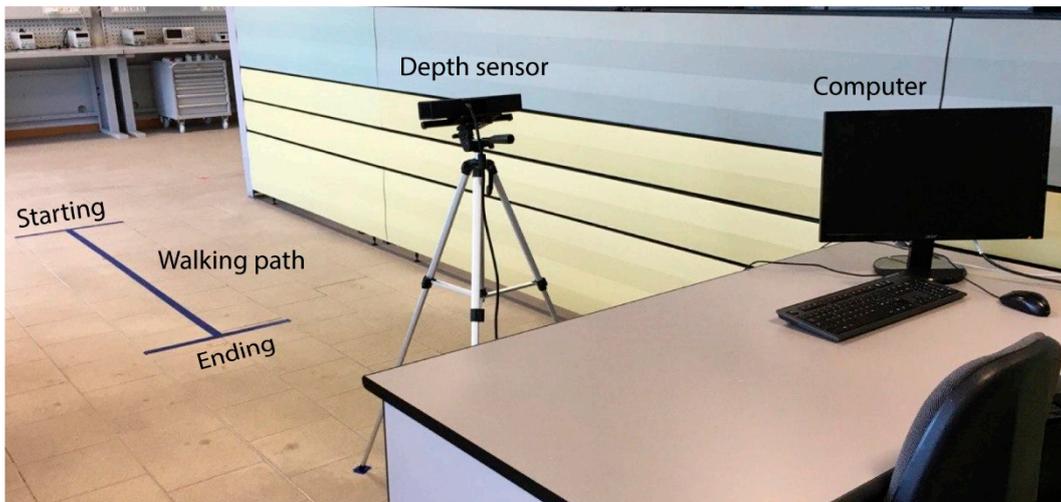


Figure 3. Experiment layout.

Table 2. Sessions, days and evaluators orders.

	First Day	Seven Days Later
Eval1	Session 1 Session 2	Session 4
Eval2	Session 3	

The measures obtained from Session 1 and Session 2 (both performed by Eval 1 on the first day) were used to estimate the intra-rater reliability. Session 1 (performed by Eval 1) and Session 3 (performed by Eval 2) were used to estimate the inter-rater reliability. Session 2 (performed by Eval 1 on the first day) and Session 4 (performed by Eval 1 on seven days later) were used to estimate the test-retest reliability [12] (Table 3). Since the evaluators must give instructions to the subjects during the sessions, intra-rater reliability and inter-rater reliability allow to estimate the influence of the evaluators and the standardization of the task in the acquired measures when they are collected repeatedly by an evaluator (intra-rater reliability) and when they are collected by different evaluators (inter-rater reliability). On the other hand, since the subject behavior influences the walking performance, test-retest reliability allows to estimate the influence of the subject behavior in the acquired measures when the same subject perform the same task two or more times.

Table 3. Sessions used for reliability estimation.

	Intra-Rater Reliability	Inter-Rater Reliability	Test-Retest Reliability
Eval 1	Session 1 (1st day) Session 2 (1st day)	Session 1 (1st day)	Session 2 (1st day) Session 4 (7 days later)
Eval 2		Session 3 (1st day)	

2.4. Statistical Analysis

In order to measure the test-retest reliability, the inter-rater reliability and the intra-rater reliability, the intra-class correlation coefficient (ICC) was computed with 95% confidence interval (CI). Table 4 shows the value ranges used to interpret ICC values [28]. ICC for test-retest reliability and intra-rater reliability was based on 2-way mixed effects, using mean of multiple measurements and absolute agreement, according to advice for choosing ICC in [29–32]. The intra-rater reliability allowed to assess the agreement between repeated measures obtained by one evaluator that used SANE to test the same group of subjects. The test-retest reliability allowed to assess the agreement between measures obtained at different times (7-day interval) by one evaluator that use SANE to the same group of subjects. ICC for inter-rater reliability was based on two-way random effects using average measures with absolute agreement, according to advice for choosing ICC in [29,31], so that the systematic errors of both evaluators (Eval 1 and Eval 2) and random residual errors were taken into account. The inter-rater reliability allows to assess the agreement between the measures obtained by Eval 1 and Eval 2 that used SANE to test the same group of subjects.

Table 4. Intra-class correlation coefficient (ICC) values interpretation [28].

ICC	Interpretation
0–0.39	Poor reliability
0.4–0.74	Modest reliability
0.75–1	Excellent reliability

In addition, Standard Error of Measurement (SEM) and Smallest Detectable Change (SDC) were calculated. SEM was calculated used the formula [33]:

$$SEM = SD \sqrt{1 - ICC} \tag{1}$$

where SD corresponds to the Standard Deviation of measurements. To calculate SEM% (independent of the measurements units), the following formula was used [34]:

$$SEM\% = (SEM/\bar{X})100 \tag{2}$$

where \bar{X} is the mean for all measurements from the evaluated sessions. SDC was calculated as [35]:

$$SDC = 1.96 \sqrt{2} SEM \tag{3}$$

To calculate SDC% (independent of the measurements units), the following formula was used [34]:

$$SDC\% = (SDC/\bar{X})100 \tag{4}$$

SEM% < 5% is considered acceptable in this study. Score interpretation of SDC% is shown in Table 5 [25,36]. All statistical calculations were carried out in Matlab 2017b (MathWorks Inc., Natick, MA, USA); ICC algorithms in [37,38] were used.

Table 5. Smallest Detectable Change (SDC) values interpretation [25,36].

SDC%	Interpretation
<10%	Excellent
10–30%	Acceptable
>30%	Poor

3. Results

3.1. Gait Cycle Duration

Results for gait cycle duration are shown in Table 6. Measurements of gait cycle duration obtained an excellent intra-rater reliability (ICC = 0.87) with an acceptable SEM (2.7%) and an excellent SDC (7.6%). The test-retest reliability was modest (ICC = 0.62) with a SEM of 5.5% and an acceptable SDC (15.2%). The inter-rater reliability was modest (ICC = 0.72) with a SEM of 5.9% and an acceptable SDC (16.4%).

Table 6. Results for gait cycle duration.

		Mean ± SD (s)	ICC	SEM	SDC
Intra-rater reliability	Session 1	1.13 ± 0.20	0.87 (0.47–0.96)	0.06 (2.7%)	0.17 (7.6%)
	Session 2	1.20 ± 0.15			
Test-retest reliability	Session 2	1.20 ± 0.15	0.62 (−0.34–0.89)	0.13 (5.5%)	0.37 (15.2%)
	Session 4	1.24 ± 0.29			
Inter-rater reliability	Session 1	1.13 ± 0.20	0.72 (0.11–0.92)	0.14 (5.9%)	0.38 (16.4%)
	Session 3	1.22 ± 0.33			

3.2. Gait Step Duration

Results for gait step duration are shown in Table 7. Measurements of gait step duration got an excellent intra-rater reliability (ICC = 0.87) with an acceptable SEM (2.7%) and an excellent SDC (7.6%). The test-retest reliability was modest (ICC = 0.62) with a SEM of 5.5% and an acceptable SDC (15.2%). The inter-rater reliability was modest (ICC = 0.72) with a SEM of 5.9% and an acceptable SDC (16.4%).

Table 7. Results for gait step duration.

		Mean ± SD (s)	ICC	SEM	SDC
Intra-rater reliability	Session 1	0.56 ± 0.10	0.87 (0.47–0.96)	0.03 (2.7%)	0.09 (7.6%)
	Session 2	0.60 ± 0.07			
Test-retest reliability	Session 2	0.60 ± 0.07	0.62 (−0.34–0.89)	0.07 (5.5%)	0.19 (15.2%)
	Session 4	0.62 ± 0.15			
Inter-rater reliability	Session 1	0.56 ± 0.10	0.72 (0.11–0.92)	0.07 (5.9%)	0.19 (16.4%)
	Session 3	0.61 ± 0.16			

3.3. Gait Cadence

Results for gait cadence are shown in Table 8. Measurements of gait cadence presented an excellent intra-rater reliability (ICC = 0.80), with an acceptable SEM (3.6%) and an excellent SDC (9.9%). The test-retest reliability was excellent (ICC = 0.77), with an acceptable SEM (3.5%) and an excellent SDC (9.7%). The inter-rater reliability was excellent (ICC = 0.87), with an acceptable SEM (4.1%) and an acceptable SDC (11.4%).

Table 8. Results for gait cadence.

		Mean ± SD (Step/min)	ICC	SEM	SDC
Intra-rater reliability	Session 1	110.12 ± 22.04	0.80 (0.26–0.94)	7.56 (3.6%)	20.94 (9.9%)
	Session 2	101.69 ± 11.94			
Test-retest reliability	Session 2	101.69 ± 11.94	0.77 (0.17–0.93)	7.08 (3.5%)	19.61 (9.7%)
	Session 4	100.01 ± 17.37			
Inter-rater reliability	Session 1	110.12 ± 22.04	0.87 (0.57–0.96)	8.80 (4.1%)	24.38 (11.4%)
	Session 3	104.40 ± 26.80			

3.4. Gait Cycle Length

Results for gait cycle length are shown in Table 9. Measurements of gait cycle length presented a modest intra-rater reliability (ICC = 0.74), with an acceptable SEM (2.4%) and an excellent SDC (6.6%). The test-retest reliability was modest (ICC = 0.68), with an acceptable SEM (2.7%) and an excellent SDC (7.6%). The inter-rater reliability was modest (ICC = 0.70), with an acceptable SEM (2.9%) and an excellent SDC (8.1%).

Table 9. Results for gait cycle length.

		Mean ± SD (m)	ICC	SEM	SDC
Intra-rater reliability	Session 1	0.95 ± 0.10	0.74 (0.03–0.92)	0.05 (2.4%)	0.13 (6.6%)
	Session 2	1.01 ± 0.08			
Test-retest reliability	Session 2	1.01 ± 0.08	0.68 (−0.02–0.91)	0.06 (2.7%)	0.16 (7.6%)
	Session 4	1.06 ± 0.12			
Inter-rater reliability	Session 1	0.95 ± 0.10	0.70 (0.07–0.91)	0.06 (2.9%)	0.16 (8.1%)
	Session 3	1.00 ± 0.11			

3.5. Gait Mean Velocity

Results for mean velocity measurements are shown in Table 10. Gait mean velocity measurements obtained an excellent intra-rater reliability (ICC = 0.95), with an acceptable SEM (2.1%) and an excellent SDC (5.8%). The test-retest reliability was excellent (ICC = 0.81), with an acceptable SEM (3.3%) and an excellent SDC (9.2%). The inter-rater reliability was excellent (ICC = 0.95), with an acceptable SEM (2.0%) and an excellent SDC (5.4%).

Table 10. Results of gait mean velocity.

		Mean ± SD (m/s)	ICC	SEM	SDC
Intra-rater reliability	Session 1	0.86 ± 0.14	0.92 (0.74–0.98)	0.04 (2.1%)	0.10 (5.8%)
	Session 2	0.85 ± 0.12			
Test-retest reliability	Session 2	0.85 ± 0.12	0.81 (0.35–0.95)	0.06 (3.3%)	0.16 (9.2%)
	Session 4	0.87 ± 0.15			
Inter-rater reliability	Session 1	0.86 ± 0.14	0.95 (0.82–0.98)	0.03 (2.0%)	0.09 (5.4%)
	Session 3	0.88 ± 0.15			

4. Discussion

This study aimed to assess the test-retest, inter-rater and intra-rater reliability of SANE for estimating gait parameters: cycle duration, step duration, cadence, cycle length and mean velocity. The test-retest, inter-rater and intra-rater reliability were excellent for gait mean velocity. The intra-rater reliability was excellent for gait cycle duration, step duration, cadence and mean velocity, and modest for gait cycle length. The test-retest was excellent for gait cadence and mean velocity, and modest for gait

cycle duration, step duration and cycle length. The gait mean velocity showed an excellent inter-rater reliability, and the other valued parameter presented a modest inter-rater reliability. No parameters presented a poor correlation. Summarizing for all parameters, test-retest reliability ranged from 0.62 to 0.81, inter-rater reliability ranged from 0.70 to 0.95 and intra-rater ranged from 0.74 to 0.92.

Since inter-rater and intra-rater reliability were close with small differences for all parameters, the training of the evaluators and the standardization of the task indicate to have a small effect on the reliability of SANE. On the other hand, test-retest reliability was lower than intra-rater and inter-rater reliability showing that the subject behavior has a greater effect on the reliability of SANE than the evaluator performance and the standardization of the task. Although it was suggested to the subjects to carry out a walk as natural as possible, it is evident that it is difficult to perform an identical walk seven days later.

The authors found that reliability studies for spatiotemporal parameters of the gait using the depth sensors Kinect usually focus on inter-method reliability comparing their measurements with the measurements of other existing systems and few of them are related with test-retest, inter-rater and intra-rater reliability of the system itself. Therefore, the results found in the study are difficult to compare with other depth sensor-based systems in terms of test-retest, inter-rater and intra-rater reliability. In [24,33] inter-day reliability (test-retest) was calculated for depth sensor-based system (Kinect v2). In [33], a one-day interval was used to perform a second session to test gait parameters of 10 children with cerebral palsy. However, the parameters evaluated are different from those evaluated in this study thus the reliability results cannot be compared. In [24], inter-day reliability (test-retest) of spatiotemporal and kinematic variables of the gait were estimated for a depth sensor-based system (Kinect v2) using seven-day interval (with injury-free/young subjects); reliability of spatiotemporal ranged from 0.55 and 0.87 (with comfortable pace) except for ground contact time (s) that had an ICC of 0.03. Test-retest reliability of SANE is within the reliability range found in [24], since the ICCs for SANE measurements ranged from 0.62 to 0.81. Therefore, it can be said that SANE has an acceptable behavior in terms of reliability when used in different days (seven-day interval). However, strategies could be tested to help the subjects have additional references during a walk test, apart from the signs on the floor, with the aim of reducing the influence of subject behavior in the test-retest measurements. For example, music background or sound references could help set a similar pace from session to session.

In [25], inter and intra-rater reliability, standard error of measurement and minimal detectable change of gait parameters were calculated for a depth sensor-based system (Kinect v2). In [25], the intra-rater reliability for spatiotemporal parameters ranged approximately from 0.77 to 0.98 and the inter-rater reliability from 0.64 to 0.98. The sessions for the intra-rater reliability were carried out within 24 h. The intra-rater reliability of SANE ranged from 0.74 to 0.92 being comparable to the intra-rater reliability values found in [25]. In addition, the inter-rater reliability of SANE ranged from 0.70 to 0.95 being within the range of inter-rater reliability found in [25]. Therefore, it can be considered that SANE has an acceptable behavior in terms of reliability when it is used during repeated measurements by a same evaluator and when it is used by different evaluators. The authors in [25] consider that their results about inter and intra-rater reliability are supported by the preliminary studies in [24,26]. However, in [24] (reference already addressed in the previous paragraph) the results can be better compared with the test-retest reliability type since the reliability in [24] was estimated using seven days interval between sessions. In [26], inter-trial reliability (which can be considered intra-rater reliability) for a depth sensor-based system (Kinect v2) ranged from 0.73 to 0.96 for spatiotemporal parameters. Therefore, intra-rater reliability of SANE and the Kinect based system in [25] are within the range of the results found in [26]. In [39], inter and intra-rater reliability were assessed for a depth sensor-based system (Kinect) for subject standing height, vertical jump height and broad jump length, finding that the reliability is comparable to clinically accepted manual measurements. However, in [39], spatiotemporal parameter of the gait was not addressed.

Test-retest reliability (from 0.62 to 0.81), inter-rater reliability (from 0.70 to 0.95) and intra-rater reliability (from 0.74 to 0.92) of SANE are comparable with those of the one of the most popular gait analysis systems VICON (three-dimensional motion analysis system): in [40], test-retest reliability and inter-tester reliability of VICON 140™ ranged from 0.97 to 0.99 for the sagittal plane, from 0.76 to 0.98 for the frontal plane and from 0.73 to 0.90 for the transverse plane; in [41], inter-session (between two sessions with a two-day interval) was >0.82 and the intra-session (between five trials) was ≥ 0.78 for VICON system. Furthermore, the test-retest reliability for spatiotemporal parameters of the three-dimensional motion analysis system (3DMA) (not Kinect) used in [24] ranged from 0.63 to 0.85 except for ground contact time(s) that had an ICC of 0.21. Therefore, test-retest reliability of SANE is comparable with 3DMA. On the other hand, it is important to mention that the authors in [25] found that the inter-rater and intra-rater reliability obtained by their Kinect v2-based system are comparable to the clinical tools 10MWT (10 Meter Walk Test, used to assess walking speed) and DGI (Dynamic Gait Index, to assess gait, balance and fall risk). As well as the authors in [26] demonstrated that their Kinect v2-based system showed a high inter-method reliability for spatiotemporal parameters when it is compared with the system GAITRite mat. In addition, it is important to mention that SANE showed an excellent inter-method reliability when it was compared with the measures of VICON cameras system in a previous study in [20].

The SEM% for reliability tests for gait cadence, cycle length and mean velocity were acceptable ($<5.0\%$). Although SEM% of test-retest and inter-rater reliability for gait cycle duration and step duration were $\leq 5.9\%$ (close to accepted values), SEM% values of SANE are comparable to the results for spatiotemporal variables in [24] where SEM% using Kinect v2 was $<16.8\%$ and using 3DMA system $<15.2\%$. SDC% for all reliability tests were from acceptable to excellent ($<16.4\%$). SDC% values of SANE are comparable to the results in [25] for spatiotemporal variables using Kinect v2 where SDC% was $<23.9\%$, except for Double support time and Step asymmetry.

5. Conclusions

Although this study has some limitations (gait analysis was obtained from 12 subjects) the results are promising. Future studies intend to expand the sample as well as evaluating SANE as measuring instrument to assess spatiotemporal gait parameters in subjects with gait impairments. SANE showed from an acceptable to an excellent test-retest, inter-rater and intra-rater reliability; test-retest reliability ranged from 0.62 to 0.81, inter-rater reliability ranged from 0.70 to 0.95 and intra-rater ranged from 0.74 to 0.92. The reliability values of SANE were comparable with other Kinect-based studies and with the most popular gait analysis systems. Since subject behavior had a greater effect on the reliability of SANE than the evaluator performance, in future work, the effects of the subjects' dresses on the intra-rater and test-retest reliability of SANE could be studied, as well as the effects of using reference sounds or music background during the test. SANE is a complete, non-intrusive and low-cost system to assess spatiotemporal gait parameters with great potential for assessing spatiotemporal parameters. The main advantage of SANE is that it automatically detects each gait cycle and automatically calculates spatiotemporal gait parameters. Since few studies address test-retest, inter-rater and intra-rater reliability of Kinect-based systems, the authors consider that the found results in this paper can be useful for other researchers. The results in this study can serve as a baseline to compare other Kinect-based systems in terms of the agreement between measures related to repetitions, time and evaluators.

Author Contributions: Conceptualization, B.D.M.C.-R. and D.C.; methodology, B.D.M.C.-R. and D.C.; software, D.C.; validation, B.D.M.C.-R. and D.C.; formal analysis, B.D.M.C.-R. and D.C.; investigation, B.D.M.C.-R. and D.C. resources, B.D.M.C.-R. and D.C.; data curation, B.D.M.C.-R.; writing—original draft preparation, B.D.M.C.-R. and D.C.; writing—review and editing, B.D.M.C.-R. and D.C.; visualization, B.D.M.C.-R. and D.C.; supervision, D.C.; project administration, D.C. All authors have read and agreed to the published version of the manuscript.

Funding: This work was funded by a grant from Ministero della Salute (Ricerca Corrente 2020).

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Springer, S.; Yogev Seligmann, G. Validity of the kinect for gait assessment: A focused review. *Sensors* **2016**, *16*, 194. [[CrossRef](#)] [[PubMed](#)]
2. Barak, Y.; Wagenaar, R.C.; Holt, K.G. Gait characteristics of elderly people with a history of falls: A dynamic approach. *Phys. Ther.* **2006**, *86*, 1501–1510. [[CrossRef](#)] [[PubMed](#)]
3. Wren, T.A.; Gorton, G.E., III; Ounpuu, S.; Tucker, C.A. Efficacy of clinical gait analysis: A systematic review. *Gait Posture* **2011**, *34*, 149–153. [[CrossRef](#)] [[PubMed](#)]
4. Cimolin, V.; Galli, M. Summary measures for clinical gait analysis: A literature review. *Gait Posture* **2014**, *39*, 1005–1010. [[CrossRef](#)]
5. Middleton, A.; Fritz, S.L.; Lusardi, M. Walking speed: The functional vital sign. *J. Aging Phys. Act.* **2015**, *23*, 314–322. [[CrossRef](#)]
6. Hodgins, D. The importance of measuring human gait. *Med. Device Technol.* **2008**, *19*, 42–44.
7. Yang, C.C.; Hsu, Y.L. A review of accelerometry-based wearable motion detectors for physical activity monitoring. *Sensors* **2010**, *10*, 7772–7788. [[CrossRef](#)]
8. Yoo, J.H.; Nixon, M.S. Automated markerless analysis of human gait motion for recognition and classification. *ETRI J.* **2011**, *33*, 259–266. [[CrossRef](#)]
9. Shotton, J.; Fitzgibbon, A.; Cook, M.; Sharp, T.; Finocchio, M.; Moore, R.; Kipman, A.; Blake, A. Real-time human pose recognition in parts from single depth images. In Proceedings of the CVPR 2011, Providence, RI, USA, 20–25 June 2011; pp. 1297–1304. [[CrossRef](#)]
10. Clark, R.A.; Pua, Y.H.; Fortin, K.; Ritchie, C.; Webster, K.E.; Denehy, L.; Bryant, A.L. Validity of the Microsoft Kinect for assessment of postural control. *Gait Posture* **2012**, *36*, 372–377. [[CrossRef](#)]
11. Bauer, S.; Seitel, A.; Hofmann, H.; Blum, T.; Wasza, J.; Balda, M.; Meinzer, H.P.; Navab, N.; Hornegger, J.; Maier-Hein, L. Real-time range imaging in health care: A survey. In *Real-Time Range Imaging in Health Care: A Survey, Time-of-Flight and Depth Imaging. Sensors Algorithms, and Applications. Lecture Notes in Computer Science*; Grzegorzec, M., Theobalt, C., Koch, R., Kolb, A., Eds.; Springer: Berlin/Heidelberg, Germany, 2013; Volume 8200, pp. 228–254. [[CrossRef](#)]
12. Pöhlmann, S.T.; Harkness, E.F.; Taylor, C.J.; Astley, S.M. Evaluation of Kinect 3D sensor for healthcare imaging. *J. Med. Biol. Eng.* **2016**, *36*, 857–870. [[CrossRef](#)]
13. Gabel, M.; Gilad-Bachrach, R.; Renshaw, E.; Schuster, A. Full body gait analysis with Kinect. In Proceedings of the 2012 Annual International Conference of the IEEE Engineering in Medicine and Biology Society, San Diego, CA, USA, 28 August–1 September 2012; pp. 1964–1967. [[CrossRef](#)]
14. Cafolla, D. A 3D Visual Tracking Method for Rehabilitation Path Planning. In *New Trends in Medical and Service Robotics. Mechanisms and Machine Science*; Carbone, G., Ceccarelli, M., Pisla, D., Eds.; Springer: Cham, Switzerland, 2019; Volume 65, pp. 264–272. [[CrossRef](#)]
15. Chaparro-Rico, B.D.M.; Cafolla, D.; Castillo-Castaneda, E.; Ceccarelli, M. Design of arm exercises for rehabilitation assistance. *J. Eng. Res.* **2020**, *8*, 203–218. [[CrossRef](#)]
16. Yao, L.; Xu, H.; Li, A. Kinect-based rehabilitation exercises system: Therapist involved approach. *Biomed. Mater. Eng.* **2014**, *24*, 2611–2618. [[CrossRef](#)] [[PubMed](#)]
17. Chaparro-Rico, B.D.M.; Cafolla, D.; Ceccarelli, M.; Castillo-Castaneda, E. Experimental characterization of NURSE, a device for arm motion guidance. *J. Healthc. Eng.* **2018**, 9303282. [[CrossRef](#)] [[PubMed](#)]
18. Zhao, W.; Feng, H.; Lun, R.; Espy, D.D.; Reinthal, M.A. A Kinect-based rehabilitation exercise monitoring and guidance system. In Proceedings of the 2014 IEEE 5th International Conference on Software Engineering and Service Science, Beijing, China, 27–29 June 2014; pp. 762–765. [[CrossRef](#)]
19. Haas, D.; Phommahavong, S.; Yu, J.; Krüger-Ziolek, S.; Möller, K.; Kretschmer, J. Kinect based physiotherapy system for home use. *Curr. Dir. Biomed. Eng.* **2015**, *1*, 180–183. [[CrossRef](#)]
20. Pavone, L.; Pscua, G.; Ricciuti, P.; Daniele, C. A kinect-based portable automatic gait analysis system: An experimental validation. *Biomed. J. Sci. Tech. Res.* **2019**, *17*, 12552–12555. [[CrossRef](#)]
21. Kurande, V.H.; Waagepetersen, R.; Toft, E.; Prasad, R. Reliability studies of diagnostic methods in Indian traditional Ayurveda medicine: An overview. *J. Ayurveda Integr. Med.* **2013**, *4*, 67–76. [[CrossRef](#)]
22. Geerse, D.J.; Coolen, B.H.; Roerdink, M. Kinematic Validation of a Multi-Kinect v2 Instrumented 10-m Walkway for Quantitative Gait Assessments. *PLoS ONE* **2015**, *10*, e0139913. [[CrossRef](#)]

23. Schmitz, A.; Ye, M.; Boggess, G.; Shapiro, R.; Yang, R.; Noehren, B. The measurement of in vivo joint angles during a squat using a single camera markerless motion capture system as compared to a marker based system. *Gait Posture* **2015**, *41*, 694–698. [CrossRef]
24. Mentiplay, B.F.; Perraton, L.G.; Bower, K.J.; Pua, Y.H.; McGaw, R.; Heywood, S.; Clark, R.A. Gait assessment using the Microsoft Xbox One Kinect: Concurrent validity and inter-day reliability of spatiotemporal and kinematic variables. *J. Biomech.* **2015**, *48*, 2166–2170. [CrossRef]
25. Latorre, J.; Colomer, C.; Alcañiz, M.; Llorens, R. Gait analysis with the Kinect v2: Normative study with healthy individuals and comprehensive study of its sensitivity, validity, and reliability in individuals with stroke. *J. Neuroeng. Rehabil.* **2019**, *16*, 97. [CrossRef]
26. Dolatabadi, E.; Taati, B.; Mihailidis, A. Concurrent validity of the Microsoft Kinect for windows v2 for measuring spatiotemporal gait parameters. *Med. Eng. Phys.* **2016**, *38*, 952–958. [CrossRef]
27. Rousson, V.; Gasser, T.; Seifert, B. Assessing intrarater, interrater and test–retest reliability of continuous measurements. *Stat. Med.* **2002**, *21*, 3431–3446. [CrossRef] [PubMed]
28. Fleiss, J.L. *Design and Analysis of Clinical Experiments*; John Wiley & Sons: Hoboken, NJ, USA, 2011; Volume 73.
29. Koo, T.K.; Li, M.Y. A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *J. Chiropr. Med.* **2016**, *15*, 155–163. [CrossRef] [PubMed]
30. Portney, L.G.; Watkins, M.P. *Foundations of Clinical Research: Applications to Practice*; Prentice Hall: Upper Saddle River, NJ, USA, 2000.
31. Perinetti, G. StaTips Part IV: Selection, interpretation and reporting of the intraclass correlation coefficient. *South Eur. J. Orthod. Dentofac. Res.* **2018**, *5*, 3–5. [CrossRef]
32. Perinetti, G. StaTips Part II: Assessment of the repeatability of measurements for continuous data. *South Eur. J. Orthod. Dentofac. Res.* **2016**, *3*, 33–34. [CrossRef]
33. Ma, Y.; Mithraratne, K.; Wilson, N.C.; Wang, X.; Ma, Y.; Zhang, Y. The validity and reliability of a Kinect v2-Based gait analysis system for children with cerebral palsy. *Sensors* **2019**, *19*, 1660. [CrossRef]
34. Nair, P.M.; Hornby, T.G.; Behrman, A.L. Minimal detectable change for spatial and temporal measurements of gait after incomplete spinal cord injury. *Top. Spinal Cord Inj. Rehabil.* **2012**, *18*, 273–281. [CrossRef]
35. Van Lummel, R.C.; Walgaard, S.; Hobert, M.A.; Maetzler, W.; Van Dieën, J.H.; Galindo-Garre, F.; Terwee, C.B. Intra-Rater, Inter-Rater and Test-Retest Reliability of an Instrumented Timed Up and Go (iTUG) Test in Patients with Parkinson’s Disease. *PLoS ONE* **2016**, *11*, e0151881. [CrossRef]
36. Llorens, R.; Latorre, J.; Noé, E.; Keshner, E.A. Posturography using the Wii Balance Board™: A feasibility study with healthy adults and adults post-stroke. *Gait Posture* **2016**, *43*, 228–232. [CrossRef]
37. Salarian, A. Intraclass Correlation Coefficient (ICC). MATLAB Central File Exchange. Available online: <https://www.mathworks.com/matlabcentral/fileexchange/22099-intra-class-correlation-coefficient-icc> (accessed on 26 May 2020).
38. Matthe, R. f_ICC. GitHub. Available online: https://www.github.com/robertpetermatthew/f_ICC (accessed on 26 May 2020).
39. Bates, N.; McPherson, A.; Berry, J.; Hewett, T. Inter- and intra-rater reliability of performance measures collected with a single-camera motion analysis system. *Int. J. Sports Phys. Ther.* **2017**, *12*, 520–526.
40. Tsushima, H.; Morris, M.E.; McGinley, J. Test-retest reliability and inter-tester reliability of kinematic data from a three-dimensional gait analysis system. *J. Jpn. Phys. Ther. Assoc.* **2003**, *6*, 9–17. [CrossRef] [PubMed]
41. Wedege, P.; Steffen, K.; Strøm, V.; Opheim, A.I. Reliability of three-dimensional kinematic gait data in adults with spinal cord injury. *J. Rehabil. Assist. Technol. Eng.* **2017**, *4*, 1–10. [CrossRef] [PubMed]

