Gait & Posture 56 (2017) 42–48

Contents lists available at ScienceDirect

Gait & Posture

journal homepage: www.elsevier.com/locate/gaitpost

Full length article

Objective assessment of movement competence in children using wearable sensors: An instrumented version of the TGMD-2 locomotor subtest

Maria Cristina Bisi[a,*], G. Pacini Panebianco[a], R. Polman[b], R. Stagni[a]

[a] Department of Electrical, Electronic and Information Engineering "Guglielmo Marconi", University of Bologna, Viale Risorgimento 2, 40136 Bologna, Italy
[b] School of Exercise and Nutrition Sciences, Queensland University of Technology, Brisbane, Australia

ARTICLE INFO

Keywords:
Children movement competence
Test instrumentation
Inertial sensors

ABSTRACT

Movement competence (MC) is defined as the development of sufficient skill to assure successful performance in different physical activities. Monitoring children MC during maturation is fundamental to detect early minor delays and define effective intervention. To this purpose, several MC assessment batteries are available. When evaluating movement strategies, with the aim of identifying specific skill components that may need improving, widespread MC assessment is limited by high time consumption for scoring and the need for trained operators to ensure reliability. This work aims to facilitate and support the assessment by designing, implementing and validating an instrumented version of the TGMD-2 locomotor subtest based on Inertial Measurement Units (IMUs) to quantify MC in children rapidly and objectively. 45 typically developing children, aged 6–10, performed the TGMD-2 locomotor subtest (six skills). During the tests, children wore five IMUs mounted on lower back, on ankles and on wrists. Sensor and video recordings of the tests were collected. Three expert evaluators performed the standard assessment of TGMD-2. Using theoretical and modelling approaches, algorithms were implemented to automatically score children tests based on IMUs' data. The automatic assessment, compared to the standard one, showed an agreement higher than 87% on average on the entire group for each skill and a reduction of time for scoring from 15 to 2 min per participant. Results support the use of IMUs for MC assessment: this approach will allow improving the usability of MC assessment, supporting objectively evaluator decisions and reducing time requirement for the evaluation of large groups.

1. Introduction

Movement competence (MC) is defined as the development of sufficient skill and ability to assure successful performance in a variety of physical activities. It has been shown an important building block in psychological, physiological, behavioural, and cognitive development of young people [1–3]. For example, being a competent mover is an important determinant of physical activity and play behaviour in young people. Children with low MC and/or with minor motor problems are more likely to be inactive, experiencing psychological and physical health problems and overall poorer well-being [4,5] and inferior cognitive development (e.g. academic performance and language) [3].

Fundamental motor skills (FMS) are at the basis of MC and failure to develop them during preschool and school years often leads to failure in the mastery of these skills during adulthood [6]. Game experiences and organized programs influence the progress in FMS [7], however explicit instruction and guidance is required to appropriately develop FMS [8]. From these premises, it is clear that monitoring children MC during maturation would allow to detect even minor delays as early as possible and develop effective interventions.

A number of different test batteries have been proposed to screen and evaluate the performance of FMS in typically developing children [9]. Product-oriented assessments evaluate the outcome of a movement (e.g. how fast, how many), offer an objective evaluation of the outcome of the task, but do not allow interpretation on how it was achieved. On the other hand, process-oriented motor competence assessments analyse how a movement is performed and with which strategy, with the advantage of allowing the identification of specific skill components that may need improving [10,11].

A particular limitation of process-oriented assessment is that it is time consuming and requires the involvement of numerous trained observers to ensure reliability. In general, the use of combined process and product assessments is suggested, if a complete and comprehensive capture of the motor development of the child is to be made [12,13]. On the other hand, the limitations described above of process-oriented tests limit the possibility of large-scale implementation for wide spread monitoring.

Among process-oriented assessments, the Test of Gross Motor

* Corresponding author.
E-mail address: mariacristina.bisi@unibo.it (M.C. Bisi).

http://dx.doi.org/10.1016/j.gaitpost.2017.04.025
Received 22 November 2016; Received in revised form 19 April 2017; Accepted 20 April 2017

Development – 2 (TGMD-2) [10] is widely used for research in several countries [14–16] with the primary purpose of identifying and screening children demonstrating delayed FMS competence. TGMD-2 evaluates children's FMS based on the presence or absence of 3–5 performance criteria for 12 skills (6 locomotor and 6 object control skills). Usually, it requires video recording of the tests and postponed evaluation by a trained expert operator [11]. Overcoming the limitations of the need of video collection, post-hoc analysis, and requirement of trained evaluators would allow wide spread administration of this test for effective screening of large populations. This can be achieved by exploiting methods used in instrumented human movement analysis. In particular, the use of wearable Inertial Measurement Units (IMUs), able to accurately quantify and consistently record human movement, has steadily risen in recent years in clinical contest, particularly for elderly and/or pathological adult populations [17].

Few studies evaluated motor development in children using IMUs. Bisi and Stagni [18] analysed the development of gait in healthy toddlers, highlighting the importance of longitudinal studies on children motor development, given their high inter-subject variability. In 2016 [19], they suggested complexity of trunk acceleration signal during gait as a possible characterizing parameter of gait maturation. Masci et al. [20,21] and Grimpampi et al. [22] examined the use IMUs and temporal and/or kinematic parameters to assess developmental levels in some FMS. In particular, they analysed hopping [20], running [21], and overarm throwing [22], suggesting this technology as a promising in-field and user-independent motor development assessment tool, for health-care, physical education and sport training professionals. According to these studies, quantitative data are essential, not only for screening purpose, but also for increasing and furthering knowledge on the mechanisms of motor development, through the longitudinal analysis of the development of strategies for achieving different skills during maturation [18].

The aim of this work was to propose a novel use of IMUs for MC batteries instrumentation. Quantitative data and ad-hoc developed algorithms can provide automatically standard scoring of MC assessment, allowing development of intervention solutions that are based on reliable and objective data, low in time consumption, easy to use and that do not require video recording.

To verify the feasibility of this approach, this first work focused on TGMD-2 locomotor subtest skill assessment. An instrumented version of TGMD-2 locomotor subtest was designed, implemented and validated as a new quantitative tool for its objective and rapid assessment based on wearable IMUs.

2. Materials & methods

2.1. Study participants

Forty-five typically developing Italian children aged 6–10 years participated in the study. Three groups of 15 children each were divided by age (Table 1).

All children had no known developmental delay and no musculoskeletal pathology.

The Bioethical Committee of the University of Bologna approved this study, and informed consent was obtained from the participants' parents.

Table 1
The content of all cells should be alligned with titles in the first raw. The last column (on the right) is empty and can be erased. Participants' characteristics for the three groups (6YC = 6 year old children, 8YC = 8 year old children, 10 YC = 10 year old children).

	age (y)	females	males	height (cm)	mass (kg)
6YC	6 ± 0	4	11	120 ± 3	23 ± 3
8YC	8 ± 0	7	8	131 ± 7	29 ± 6
10YC	10 ± 0	6	9	143 ± 8	38 ± 6

2.2. Experimental protocol

Five tri-axial wireless IMUs (OPALS, Apdm, USA) were mounted using straps on the lower back, ankles (above the lateral malleolus) and wrists.

Children were asked to wear comfortable clothes and gym shoes. They were asked to perform the TGMD-2 locomotor subtest, which consists in six tasks: run, gallop, hop, leap, horizontal jump and slide [10]. Children were instructed according to TGMD-2 guidelines. Prior to the testing, a standing static posture was acquired. Measures of accelerations, angular velocities and directions of magnetic field were collected using the IMUs (sampling frequency, 128 Hz). Videos were also recorded during the tests using two different video recorders (frontal and sagittal plane, GoProHero 4, GoPro Inc. USA, and Canon Legria FS20, Canon Europe).

Each participant performed 2 trials for each task resulting in a total of 90 trials (30 for each age category). Due to participant unfamiliarity or lack of cooperation, some tasks were not performed by all children (e.g. for the leap only 78 trials were recorded).

2.3. Data analysis: standard assessment

To ensure reliable standard assessment, three expert operators performed the standard scoring of the different skills based on video-recordings. For each skill, the TGMD-2 provides a variable number (ranging from 3 to 5) of performance criteria (*pc*) to be evaluated. As described in TGMD-2 guidelines:

1) Scores were assigned for each *pc*. The child was given 1 for a pass, 0 for a failed attempt (no partial marks).
2) The Subtest Raw Score was obtained by adding up the 6 skill scores of the locomotor subtest (high scores indicate better performance than low scores).
3) Subtest Raw Scores were converted to Subtest Standard Scores, which take into account age and sex of the child.
4) Descriptive Ratings ranging from 'Very poor' to 'Mastery' were assigned to the Subtest Standard Score of each child.

All scorings were done by analysing the videos after the whole assessment. Each rater viewed and scored the videos individually for each participant independently. When there was no uniform agreement amongst operators on the *pc* score, the score with two votes was assigned.

2.4. Data analysis: automatic assessment

Using theoretical approaches and modelling hypotheses, different algorithms were defined to score automatically each *pc*, based on the data collected by IMUs. Only Running *pc* number 3 ("Narrow foot placement, landing on heel or toe") was excluded from the analysis due to the too high variability in the soles of gym shoes worn by children that did not allow a reliable analysis of foot contact impact [23–25]. A total of 23 different algorithms for 23 *pc* were developed and implemented in Matlab 2012b (MathWorks BV, USA).

A brief description of each algorithm is reported in Table 2a together with a list of relevant acronyms (Table 2b). Except the required manual loading of the data, all the algorithms are completely automatic, including counting of correct events and verification of correct sequences of movement. Loading of the data collected during the standing static posture is required for the *Automatic Assessment* of run *pc4* and slide *pc1*. When a threshold had to be identified for the algorithm, no more than 8 children' tests (with different shown performances) were used for defining the threshold. Thresholds were designed to be generalizable when evaluating the specific task (e.g.

Table 2
(a) List of performance criteria (*pc*) and corresponding algorithms' brief description for each tasks. (b) List of acronyms

(a)		
1. Run		
pc1	Arms move in opposition to legs, elbows bent	Peaks in ωML of opposite arms and legs should coincide and have approximately the same amplitude. Threshold for arm ωML peaks was fixed at 75% of leg ωML peaks.
pc2	Brief period where both feet are off the ground	Correct sequence of foot contact and foot off events has to be present. Foot contacts and foot offs are identified on the leg ωML.
pc3	Narrow foot placement, landing on heel or toe	//
pc4	Nonsupport leg bent approximately 90 ° (i.e., close to buttocks)	Analysis of the magnetic field measured on the V axis of the support leg at Foot off. If the non support leg is correctly at 90°, the support leg at Foot off is inclined forward. Threshold has to be fixed in respect to signals obtained during the specific static postural acquisition (e.g. in our work threshold was fixed for all the participants at −50% of the magnetic field measured on the leg V axis in the static postural position).
2. Gallop		
pc1	Arms bent and lifted to waist level at takeoff	Analysis of the median acceleration components of the forearms. The identification of approximate position is based on the components of static gravity acceleration along the axes on the sagittal plane.
pc2	A step forward with the lead foot followed by a step with the trailing foot to a position adjacent to or behind the lead foot	After identifying the preferred leg, the correct sequence of foot contacts and foot offs (tFC lFC tFO lFO) is verified. Foot contacts and foot offs are identified on the leg ωML. A threshold is fixed for limiting the time distance between swing phases. Time distance between the swing phases of the lead and the trailing leg has to range between 10% and 80% of the time distance between the trailing leg and the subsequent swing of the lead leg. The inferior limit excludes the coincidence of the swing events (task performed like a jump).
pc3	Brief period where both feet are off the floor	Flight time is estimated as the time distance between toe off of the following leg and heel strike of the preferred leg.
pc4	Maintains a rhythmic pattern for four consecutive gallops	Correct galloping events are counted: the correctness is based on the above described criteria.
3. Hop		
pc1	Nonsupport leg swings forward in pendular fashion to produce force	Analysis of phases of leg ωML: the point is assigned if the support leg's phase is opposite to the non support leg's one. Amplitudes of the two ωML should be close.
pc2	Foot of nonsupport leg remains behind body	Analysis of median acceleration components of the legs: gravity components should be majorly on the V axis of the supporting leg and on the AP axis of the non support leg.
pc3	Arms flexed and swing forward to produce force	Analysis of the median acceleration components of the forearms. The identification of approximate position is based on the components of static gravity acceleration along the axes on the sagittal plane. Analysis of phases of ML angular velocities of the arms and support leg: the point is assigned if the support leg's phase is opposite to arms' one. Amplitude of ωML of arms and leg should be close.
pc4	Takes off and lands three consecutive times on preferred foot	Hopping events on preferred foot are counted. Foot contacts estimated from the preferred leg ωML.
pc5	Takes off and lands three consecutive times on other foot	Hopping events on nonpreferred foot are counted. Foot contacts estimated from the nonpreferred leg ωML.
4. Leap		
pc1	Take off on one foot and land on the opposite foot	At first, time of foot land is obtained looking for the peak of trunk accV. Foot contacts are identified on the two legs ωML: land foot is identified as the one with one foot contact close to the time of foot land, take off foot as the one with the last foot contact prior to the time of foot land. Correct sequence of alternating feet is verified.
pc2	A period where both feet are off the ground longer than running	Median run flight time is estimated from Run.pc2 results. Leap flight time is estimated as the time distance between take off and land. Criterion: Leap flight time > 1.2 x run flight time.
pc3	Forward reach with the arm opposite the lead foot	Peak of ωML of arms and legs during the task should be close both in time and in amplitude. Positive/negative signs of ωML of the leading leg and of the opposite arm are analysed in order to verify the correct forward reach.
5. Horizontal Jump		
pc1	Preparatory movement includes flexion on both knees with arms extended behind body	After identifying the instant of take off, ωML of the arms and of the legs are analyzed: ωML in the correct direction have to be present on both arms and legs prior to take off.
pc2	Arms extend forcefully forward and upward reaching full extension above the head	After identifying the instant of take off, arm ωML is analyzed: peak velocities have to be present in the period between take off and the middle of the flight.
pc3	Take off and land on both feet simultaneously	Foot off and foot landing instants are identified using the wavelet transform on the leg accAP. A threshold of 0.08 s (10 sample) is fixed for identifying simultaneous take off and land on both feet.
pc4	Arms are thrust downward during landing	After identifying the instant of foot landing, ωML of the arms is analyzed: peak velocities have to be present in the period between the middle of the flight and landing instant.
6. Slide		
pc1	Body turned sideways so shoulders are aligned with the line on the floor	Body position is obtained comparing trunk magnetometer signals during sliding with the ones obtained during the static calibration test.
pc2	A step sideways with the lead foot followed by a slide of the trailing foot to a point next to the lead foot	After identifying the front leg, correct sequence of swing periods and foot contact events is verified. Foot contacts are identified on the leg ωAP.
pc3	A minimum of four continuous step-slide cycles to the right	Correct step-sliding cycles to the right are counted: the correctness is based on the above described criteria.
pc4	A minimum of four continuous step-slide cycles to the left	Correct step-sliding cycles to the left are counted: the correctness is based on the above described criteria.

(b)

AP	antero-posterior
ML	medio-lateral
V	vertical
ωAP	angular velocity about AP axis
ωML	angular velocity about ML axis
accAP	acceleration along AP axis
accV	acceleration along V axis
tFC	trailing foot contact
lFC	lead foot contact
tFO	trailing foot off
lFO	lead foot off

ratio between peaks in leg and arms ML angular velocity).

After the automatic scoring of each *pc*, Subtest Raw Scores, Standard Scores and corresponding Descriptive Ratings were calculated with the same procedure described for the *Standard Assessment*.

For the identification of foot contact events, the method proposed by Aminian [26] was used in all the tasks, except the horizontal jump, to identify local minima before and after swing phase, which were evident. Since the above mentioned method did not allow to correctly identify foot off and foot strike in horizontal jump, a method for the investigation of transient events in biomedical signals, based on a wavelet-based energetic approach, was applied [27]. Both methods, originally defined for different purposes, were preliminary validated using recorded videos as reference. Details on which signals were analyzed for contact event identification in each task are described in Table 2a.

When the task included a sequence of repetitive movements (e.g. gallop, hop, run, and slide), the first repetition was discarded and the following four were analyzed to exclude initiation and termination phases.

Table 3
Number of tests analysed (n° tests) and number of times a performance criterion (*pc*) was achieved in the entire group of children.

	n° tests	pc 1	pc 2	pc 3	pc 4	pc 5
Run	90	42	90	–	49	
Gallop	88	0	65	65	64	
Hop	86	24	61	27	81	81
Leap	78	78	66	6		
H Jump	82	33	10	78	32	
Slide	90	71	80	75	57	

2.5. Data analysis: statistics

The percentage of agreement between the *Automatic Assessment* and the *Standard Assessment* was statistically evaluated on two levels:

(a) By comparing raw scores assigned to each item over the total of the tests performed and per age group.
(b) By comparing Subtest Standard Scores to the corresponding Descriptive Ratings of each child on the entire group and per age group.

To investigate possible limitations in using the same algorithms for different age groups, a one-way ANOVA (level of significance 5%) was performed to analyze the effect of age on percentage of agreement.

The normal distribution of the differences between scores obtained using *Automatic Assessment* and *Standard Assessment* was verified using a Shapiro Wilk test. Bland Altman plots were used to compare Subtest Raw Scores and Subtest Standard Scores obtained using *Automatic* and *Standard Assessment*.

3. Results

3.1. Standard assessment

Mean time of evaluation of all the *pc* per participant on recorded videos was 15 min (excluding time for downloading and opening videos). Approximate dimension of recorded videos for the tests of a single participant was 1.8 GB for high-resolution videos (recorded by the GoPro Hero 4) and 200 MB for the low-resolution videos (recorded by Canon Legria FS20).

The inter-rater reliability was appropriate. Maximum Raw Score mean difference between two raters was 2.0 and the largest 95% confidence interval of the mean was 4.1. For Subtest Standard Scores, the maximum mean difference was 0.8 and 95% confidence interval of

Table 4
Number of tests analysed (n° tests) and percentage of agreement between Automatic and Standard Assessments for each performance criteria (*pc#*) and each task, shown per age group and on the total of tests analysed.

Run	6YC	8YC	10YC	Total	**Gallop**	6YC	8YC	10YC	Total	**Hop**	6YC	8YC	10YC	Total
n° tests	30	30	30	90	n° tests	30	30	28	88	n° tests	28	30	28	86
pc 1	90%	80%	80%	83%	pc 1	97%	100%	100%	99%	pc 1	93%	77%	82%	84%
pc 2	100%	100%	100%	100%	pc 2	83%	77%	100%	86%	pc 2	86%	83%	79%	83%
pc 3	//	//	//		pc 3	83%	77%	100%	86%	pc 3	89%	100%	96%	95%
pc 4	87%	93%	83%	88%	pc 4	83%	77%	96%	85%	pc 4	100%	100%	100%	100%
										pc 5	100%	93%	100%	98%
mean	92%	91%	88%	90%	mean	87%	83%	99%	89%	mean	94%	91%	91%	92%
Leap	6YC	8YC	10YC	Total	**H Jump**	6YC	8YC	10YC	Total	**Slide**	6YC	8YC	10YC	Total
n° tests	20	28	30	78	n° tests	24	30	28	82	n° tests	30	30	30	90
pc 1	90%	100%	97%	96%	pc 1	79%	93%	89%	88%	pc 1	97%	93%	83%	91%
pc 2	95%	82%	87%	87%	pc 2	96%	87%	93%	91%	pc 2	97%	90%	97%	94%
pc 3	80%	100%	90%	91%	pc 3	83%	77%	86%	82%	pc 3	100%	97%	100%	99%
					pc 4	88%	87%	89%	88%	pc 4	100%	100%	100%	100%
mean	88%	94%	91%	91%	mean	86%	86%	89%	87%	mean	98%	95%	95%	96%

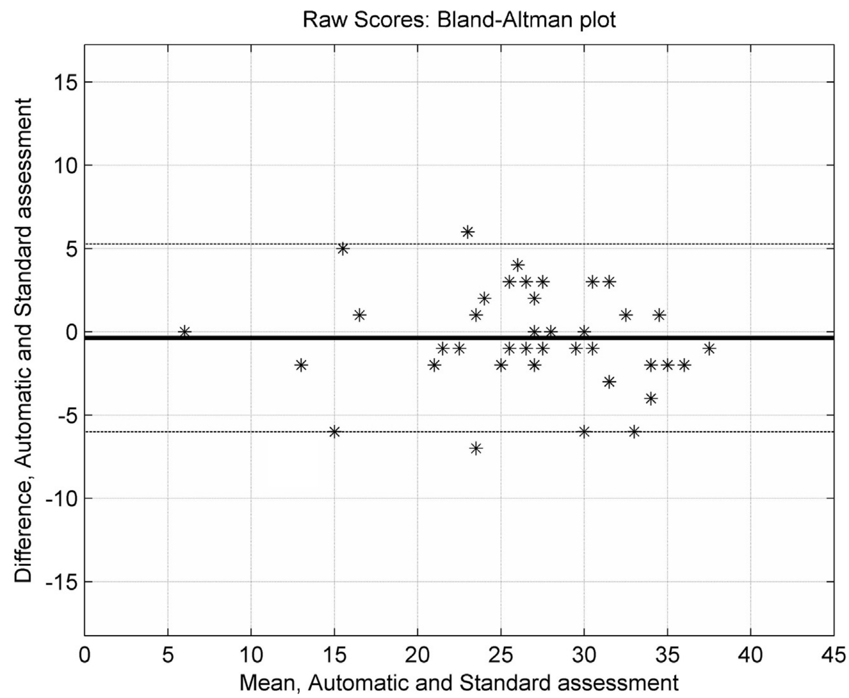

Fig. 1. Bland Altman plots of Subtest Raw Scores obtained by *Automatic Assessment* and *Standard Assessment* (mean, solid line and 95% confidence interval, dotted lines).

the mean was 1.7.

Scores assigned using *Standard Assessment* were not uniformly distributed among children: e.g. all the analysed children achieved running *pc*2 while none achieved gallop *pc*1. Table 3 shows the sum of positive assessment per *pc* of each task on the entire group. The total of available tests (n° tests) for each task is shown in Table 3 and for each group in Table 4.

Descriptive Ratings of analysed children ranged from '*very poor*' to '*below average*': 17 children showed Standard Scores classified as *'very poor'*, 20 *'poor'*, and 8 *'below average'*.

3.2. Automatic assessment

Mean time of evaluation of all the *pc* per participant on IMUs data was 2 min of computing time (excluding time for downloading sensors data). Approximate dimension of IMUs data for a single participant was 60 MB.

Algorithm results showed agreement with Raw Scores assigned visually by operators with a mean percentage (on the entire group of children) that ranged from 82% to 100%, depending on the *pc* and on the skill. No effect of age was found in the agreement results between *Automatic* and *Standard Assessment*.

Looking at the different skills, *Automatic* and *Standard Assessments* showed the best accordance on the slide, with a 96% agreement on average (min − max range 91–100%). The lowest correspondence was found for the horizontal jump with an average of 87% of agreement (min − max range 82–91%). All the algorithms showed a minimum of 77% of agreement with the corresponding performance criterion assessed with the *Standard Assessments* both on the different age groups and on the entire group (details in Table 4).

Descriptive Ratings of analyzed children obtained using *Automatic Assessment* showed an agreement of 73% with the ones obtained using *Standard Assessments*; 12 children showed differences in the descriptive ratings using the two methods. For all the 12 children, differences were within one level of evaluation; 7 children showed a higher evaluation with the *Standard Assessment* and 5 with the *Automatic* one. Bland Altman plots for Subtest Raw Scores are shown in Fig. 1. Mean difference was −0.37 and 95% confidence interval of the mean was 5.6. For Subtest Standard Scores mean differences between *Standard* and *Automatic Assessment* was −0.4 and 95% confidence interval 1.7.

4. Discussion

In this work, the use of IMUs was introduced for the instrumentation of the TGMD-2 locomotor subtest, with the aim of providing automatic standard scoring, improving objectivity, supporting evidence-based evaluators' decisions, and reducing the person-time required for the evaluation.

The proposed method was validated on tests performed by 45 children aged 6–10 years, comparing scores assigned by the novel instrumented method (*Automatic Assessment*) with the ones assigned by expert operators following TGMD-2 guidelines (*Standard Assessment*).

Study results showed good agreement between *Automatic* and *Standard Assessment* and a significant time reduction when using *Automatic Assessment* (from 15 of person-time to 2 min computing time per participant).

The agreement of each algorithm with the corresponding *pc* assessed by *Standard Assessment* on the entire group of participants ranged from 82% to 100% depending on the specific analysed *pc*. No age effect was found, indicating that there is no need of defining age specific thresholds and/or different algorithms when aiming at scoring automatically the locomotor subtest of the TGMD-2 in children between 6 and 10.

According to Bland Altman plot analysis, group results of Subtest Raw Scores and Subtest Standard Scores did not differ significantly with *Standard* and *Automatic Assessment*, resulting in a comparable MC assessment. The 95% confidence interval was inferior to 6 points on a total of 48 for the Raw Scores and inferior to 2 points on 20 for Standard Scores. Mean differences between the two methods were close to 0, showing no bias. These results are similar to those reported for intra-, inter-rater reliability in the *Standard Assessment* of TGMD-2 in this study and in previous literature [11,14,28], suggesting the automatic approach as a valid instrument for MC monitoring.

In this work, 23 algorithms were developed for assessing the corresponding criteria of the TGMD-2 locomotor subtest. Their definition was based on existing literature about methods for IMUs application in sports and clinics, and on theoretical and modelling hypotheses. The exclusion from the *Automatic Assessment* of Running *pc*3, due to the too high variability caused by wearing gym shoes by participants is a limitation of the study. In order to overcome this limitation, two possible solutions are available: i) have the evaluator manually score this *pc*, or, for a complete reliable automatic assessment, ii) provide more detailed instructions about the type of shoes required for the tests (i.e. specifying no cushioned heel shoes [24]).

As shown in Table 3, Raw Scores assigned using *Standard Assessment* were not uniformly distributed among children, implying a different level of verification for the different algorithms. In particular, all and none of the children achieved *pc*2 of run and *pc*1 of gallop, respectively. *Pc*2 of run requires verifying the sequence of foot contacts and foot offs during the task. Foot contact identification procedure used in this work is a widely used and validated method for gait event detection based on the analysis of the ML angular velocity of the shanks [26], thus a high level of concordance between the two assessment methods could be expected also when the criterion is not achieved. *Pc*1 of gallop was not achieved by any of the children. In this case, a lower performance could be expected when including tests where children both achieve and not achieve the criterion. However, the algorithm was similar to the one of hop *pc*2, which showed on average a 83% of agreement with its *Standard Assessment*.

Overall, the instrumented version of the TGMD-2 locomotor subtest was proven a valid instrument for in field MC assessment, fulfilling the requirements of being objective, reliable, easy and quick to use. In addition, the proposed instrumented version does not require video recordings and trained assessors, simplifying the assessment and reducing restrictions and requirements related to privacy and ethical issues.

The results of this study represent an important step towards the use of wearable technologies and quantitative measures in MC assessment and can be considered a starting point for many future developments. Firstly, the same approach can be applied to the object-control subtest for completing the instrumentation of the TGMD-2 and it can be used for the instrumentation of its new version TGMD-3 [29]. End-user implementations of the instrumented batteries could be designed for storing detailed information about children performance that can be analysed and/or recalled when required. This would be particularly advantageous in intervention studies or in the rehabilitation context. Future research on TGMD-2 instrumentation will investigate the possibility of identifying minimal setup solutions (e.g. based on a single IMU), in order to further enhance and promote MC monitoring (a solution based on a single sensor could be implemented directly on an App for smartphone using the on-board IMU). The inclusion of synthetic parameters already proposed for the description of MC levels [20–22] and of others found promising for motor control development characterization [19,30] could provide further support for objective developmental measures.

Finally, the instrumented version of process-oriented batteries can be integrated with product-oriented evaluations, in order to implement tools that follows the most promising approach for the assessment of MC in children [12] and at the same time are rapid, objective, and easy to use.

Conflict of interest

We wish to confirm that there are no known conflicts of interest associated with this publication and there has been no significant financial support for this work that could have influenced its outcome.

Acknowledgments

Authors would like to thank children, their parents, teachers and coordinators of Istituto San Giuseppe Lugo (Lugo (RA), Italy) that allowed data acquisition. We also would like to thank Dr. Simone Ciccioli for the help during video analysis.

References

[1] S.K. Lai, S.A. Costigan, P.J. Morgan, D.R. Lubans, D.F. Stodden, J. Salmon, L.M. Barnett, Do school-based interventions focusing on physical activity, fitness, or fundamental movement skill competency produce a sustained impact in these outcomes in children and adolescents? A systematic review of follow-up studies, Sports Med. Auckl. N. Z. 44 (2014) 67–79, http://dx.doi.org/10.1007/s40279-013-0099-9.

[2] R. Bailey*, Evaluating the relationship between physical education, sport and social inclusion, Educ. Rev. 57 (2005) 71–90, http://dx.doi.org/10.1080/0013191042000274196.

[3] M.T. Kantomaa, E. Stamatakis, A. Kankaanpää, M. Kaakinen, A. Rodriguez, A. Taanila, T. Ahonen, M.-R. Järvelin, T. Tammelin, Physical activity and obesity mediate the association between childhood motor function and adolescents' academic achievement, Proc. Natl. Acad. Sci. U. S. A. 110 (2013) 1917–1922, http://dx.doi.org/10.1073/pnas.1214574110.

[4] L.M. Barnett, E. van Beurden, P.J. Morgan, L.O. Brooks, J.R. Beard, Childhood motor skill proficiency as a predictor of adolescent physical activity, J. Adolesc. Health 44 (2009) 252–259, http://dx.doi.org/10.1016/j.jadohealth.2008.07.004.

[5] D.R. Lubans, P.J. Morgan, D.P. Cliff, L.M. Barnett, A.D. Okely, Fundamental movement skills in children and adolescents: review of associated health benefits, Sports Med. Auckl. N. Z. 40 (2010) 1019–1035, http://dx.doi.org/10.2165/11536850-000000000-00000.

[6] R. Vameghi, A. Shams, P. Shamsipour Dehkordi, The effect of age, sex and obesity on fundamental motor skills among 4 to 6 years-old children, Pak. J. Med. Sci. 29 (2013) 586–589.

[7] L.L. Hardy, T. Reinten-Reynolds, P. Espinel, A. Zask, A.D. Okely, Prevalence and correlates of low fundamental movement skill competency in children, Pediatrics 130 (2012) e390–398, http://dx.doi.org/10.1542/peds.2012-0345.

[8] S.W. Logan, L.E. Robinson, A.E. Wilson, W.A. Lucas, Getting the fundamentals of movement: a meta-analysis of the effectiveness of motor skill interventions in children, Child Care Health Dev. 38 (2012) 305–315, http://dx.doi.org/10.1111/j.1365-2214.2011.01307.x.

[9] A.W. Burton, D.E. Miller, Movement Skill Assessment, Human Kinetics, 1998.

[10] D.A. Ulrich, TGMD–2: Test of Gross Motor Development, second edition, Pro-Ed, Austin, Tx, 2000.

[11] L.M. Barnett, C. Minto, N. Lander, L.L. Hardy, Interrater reliability assessment using the test of gross motor development-2, J. Sci. Med. Sport. 17 (2014) 667–670, http://dx.doi.org/10.1016/j.jsams.2013.09.013.

[12] S.W. Logan, L.M. Barnett, J.D. Goodway, D.F. Stodden, Comparison of performance on process- and product-oriented assessments of fundamental motor skills across childhood, J. Sports Sci. (2016) 1–8, http://dx.doi.org/10.1080/02640414.2016.1183803.

[13] H.G. Williams, E.V. Monsma, Assessment of gross motor development, in: Psychoeduc. Assess. Presch. Child., IV, Mahwah, N.J.: Lawrence Erlbaum Associates, c2007., n.d.: p. 448.

[14] N.C. Valentini, Validity and reliability of the TGMD-2 for Brazilian children, J. Mot. Behav. 44 (2012) 275–280, http://dx.doi.org/10.1080/00222895.2012.700967.

[15] J. Rudd, M.L. Butson, L. Barnett, D. Farrow, J. Berry, E. Borkoles, R. Polman, A holistic measurement model of movement competency in children, J. Sports Sci. 34 (2016) 477–485, http://dx.doi.org/10.1080/02640414.2015.1061202.

[16] J. Simons, D. Daly, F. Theodorou, C. Caron, J. Simons, E. Andoniadou, Validity and reliability of the TGMD-2 in 7–10-year-old Flemish children with intellectual disability, Adapt. Phys. Act. Q. APAQ 25 (2008) 71–82.

[17] A. Godfrey, J. Lara, C.A. Munro, C. Wiuff, S.A. Chowdhury, S. Del Din, A. Hickey, J.C. Mathers, L. Rochester, Instrumented assessment of test battery for physical capability using an accelerometer: a feasibility study, Physiol. Meas. 36 (2015) N71–83, http://dx.doi.org/10.1088/0967-3334/36/5/N71.

[18] M.C. Bisi, R. Stagni, Evaluation of toddler different strategies during the first six-months of independent walking: a longitudinal study, Gait Posture 41 (2015) 574–579, http://dx.doi.org/10.1016/j.gaitpost.2014.11.017.

[19] M.C. Bisi, R. Stagni, Complexity of human gait pattern at different ages assessed using multiscale entropy: from development to decline, Gait Posture 47 (2016) 37–42, http://dx.doi.org/10.1016/j.gaitpost.2016.04.001.

[20] I. Masci, G. Vannozzi, N. Getchell, A. Cappozzo, Assessing hopping developmental level in childhood using wearable inertial sensor devices, Motor Control 16 (2012) 317–328.

[21] I. Masci, G. Vannozzi, E. Bergamini, C. Pesce, N. Getchell, A. Cappozzo, Assessing locomotor skills development in childhood using wearable inertial sensor devices: the running paradigm, Gait Posture 37 (2013) 570–574, http://dx.doi.org/10.1016/j.gaitpost.2012.09.017.

[22] E. Grimpampi, I. Masci, C. Pesce, G. Vannozzi, Quantitative assessment of developmental levels in overarm throwing using wearable inertial sensing technology, J. Sports Sci. 34 (2016) 1759–1765, http://dx.doi.org/10.1080/02640414.2015.1137341.

[23] D.E. Lieberman, M. Venkadesan, W.A. Werbel, A.I. Daoud, S. D'Andrea, I.S. Davis, R.O. Mang'Eni, Y. Pitsiladis, Foot strike patterns and collision forces in habitually barefoot versus shod runners, Nature 463 (2010) 531–535, http://dx.doi.org/10.1038/nature08723.

[24] Running Barefoot Heel Striking & Running Shoes, (n.d.). http://www.barefootrunning.fas.harvard.edu/2FootStrikes & RunningShoes.html. (Accessed 18 October 2016).

[25] S. Mullen, E.B. Toby, Adolescent runners: the effect of training shoes on running kinematics, J. Pediatr. Orthop. 33 (2013) 453–457, http://dx.doi.org/10.1097/BPO.0b013e31829241dc.

[26] K. Aminian, B. Najafi, C. Büla, P.-F. Leyvraz, P. Robert, Spatio-temporal parameters of gait measured by an ambulatory system using miniature gyroscopes, J. Biomech. 35 (2002) 689–699.

[27] E. Magosso, M. Ursino, A. Zaniboni, E. Gardella, A wavelet-based energetic approach for the analysis of biomedical signals: application to the electroencephalogram and electro-oculogram, Appl. Math. Comput. 207 (2009) 42–62, http://dx.doi.org/10.1016/j.amc.2007.10.069.

[28] M. Cano-Cappellacci, F.A. Leyton, J.D. Carreño, Content validity and reliability of test of gross motor development in Chilean children, Rev. Saúde Pública 49 (2016), http://dx.doi.org/10.1590/s0034-8910.2015049005724.

[29] D.A. Ulrich, The test of gross motor development – 3 (TGMD-3): administration, scoring, and international norms, Spor Bilim. Derg. 24 (2013) 27–33.

[30] M.C. Bisi, R. Stagni, Development of gait motor control: what happens after a sudden increase in height during adolescence? Biomed Eng. Online 15 (2016) 47, http://dx.doi.org/10.1186/s12938-016-0159-0.