Balance and strength assessment of Special Olympics athletes: How feasible and reliable is the FUNfitness test battery?

Debbie Van Biesen [1*], **Roi Charles Pineda** [2]

Abstract: This study examined the test-retest reliability and feasibility of four muscle strength and three balance tests included in the Special Olympics (SO) FUNfitness test battery. The test is used worldwide to assess physical fitness of SO athletes with intellectual disabilities (ID). A sample of 36 Belgian participants with ID (22 men, 14 women) aged 8–30 years, completed a battery of seven tests twice within a two-week time interval. We assessed test-retest reliability by means of intraclass correlation coefficients (ICC), standard error of measurement (SEM), and Bland-Altman plots. All tests demonstrated good feasibility and relative and absolute reliability. The ICC ranged between 0.75 and 0.89. All SEM values demonstrated acceptable measurement precision (SEM<SD/2). The scatter around the Bland-Altman plots were randomly distributed. Despite the promising findings, further research is recommended to determine whether these balance and strength tests are also reliable in less standardized environments such as the SO testing-area.

Introduction

Intellectual disability (ID) as per definition is characterized by significant limitations in intellectual functioning (IQ ≤ 75), and deficits in the social, conceptual and practical domain (AAIDD, 2010). Furthermore, it is well documented that many individuals with ID also have lower cardiorespiratory fitness levels compared to peers without ID (Lotan, Isakov, Kessel, & Merrick, 2004) and higher rates of obesity (Foley, Lloyd, Vogl, & Temple, 2014). Hence, these individuals are a particularly vulnerable group, with a greater risk for developing chronic health disorders associated with a sedentary lifestyle, including cardiovascular disease, type 2 diabetes and some cancers (Bandini, Curtin, Hamad, Tybor, & Must, 2005; de Rezende, Lopes, Rey-López, Matsudo, & do Carmo Luiz, 2014; Lin, Yen, Li, & Wu, 2005; Taylor, Jones, Williams, & Goulding, 2002).

Physical fitness measures can be categorized into two main domains: 1) health-related and 2) skill-related components. Health-related fitness components are important to improve physical health and include cardiorespiratory capacity, muscular strength, muscular endurance, flexibility, and body composition. Apart from the abovementioned limitations in cardiorespiratory fitness and altered body composition, persons with ID have a lower level of muscle strength when compared to people without ID (Blomqvist, Olsson, Wallin, Wester, & Rehn, 2013; Wuang, Chang, Wang, & Lin, 2013). Borji, Zghal, Zarrouk, Sahli and Rebai (2014) suggested that the lower muscle strength seen in individuals with ID is not only related to external factors, such as an inactive lifestyle, but probably also related to a central nervous system failure to activate motor units and to some abnormal intrinsic muscle properties.

Aside from health-related components of physical fitness, an individual also requires well-developed skill-related physical fitness components, such as coordination, balance, speed, agility, reaction time and power, to be able to perform activities of daily living. It is a common finding in the literature that individuals with ID have poor balance control (Enkelaar, Smulders, Van Schrojenstein

Lantman - De Valk, Geurts, & Weerdesteyn, 2012). Compared to peers without ID, individuals with ID demonstrate greater instability during both quiet standing and walking as indicated by a larger and more variable body sway, and a more laterally orientated sway pattern. Whether their balance impairments should be attributed to extrinsic causes (e.g., their lower level of physical activity), intrinsic causes (e.g., the inadequate development of the central nervous system), or to a combination of both remains unclear. Regardless of the cause, balance impairment in people with ID is related to limitations in functional capacity and increased risk of falling (Hale, Bray, & Littmann, 2007; Hsieh, Heller, & Miller, 2001; Hsieh, Rimmer, & Heller, 2012; Hsu, 2016; Lee, Lee, & Song, 2016; Sherrard, Tonge, & Ozanne-Smith, 2001).

Assessing the strength and balance of individuals with ID require careful consideration of their decreased intrinsic motivation to perform with maximal effort, and increased need for clear and understandable instructions (Hutzler & Korsensky, 2010; Schützwohl et al., 2016). Important work has been done in the nineties by Winnick and Short (1999) who developed the Brockport Physical fitness test, as a health-related criterion referenced test to assess fitness in people with various types of disabilities, including ID. More recent studies demonstrated adequate feasibility and test-retest reliability of functional fitness tests in people with Down Syndrome (Boer & Moss, 2016) and in elderly people with ID (Hilgenkamp, van Wijck, & Evenhuis, 2012). Nevertheless, a valid and reliable fitness test battery specifically feasible for all people with ID is still lacking. Additionally, although laboratory tests are the most accurate option for measuring components of physical fitness and most often considered as golden standard in terms of validity and reliability, the choice for field tests is often made because they are more practical and cost-effective (e.g., less time-consuming, less need for qualified personnel, less expensive equipment) (Fjortoft, Pedersen, Sigmundsson, & Vereijken, 2011).

Large-scale field-testing of physical fitness in people with ID, takes place yearly in connection with Special Olympics (SO), which is an international sports organization for children and adults with ID. SO offers training and competition opportunities for more than 4.2 million athletes in more than 170 countries. Besides training and competition, SO also developed diverse healthcare programs aimed to screen the health status of athletes. One of these programs is FUNfitness, one of the seven pillars within the Healthy Athletes program. FUNfitness provides screenings to evaluate the athlete's physical status and offer them one-on-one education on practical suggestions to improve their health condition and avoid injuries.

The large and unique datasets regarding physical fitness of SO athletes worldwide offer unique opportunities to develop guidelines and recommendations for researchers, and sport and health practitioners and policymakers. Up until now, however, an examination of the psychometric properties of the FUNfitness test battery is lacking. Thus, the overall aim of this pilot study is to investigate the feasibility and test-retest reliability of the SO FUNfitness test, with a focus on the balance and muscle strength tests. The first aim is to investigate whether all people with ID are able to understand the instructions and accordingly perform the strength and balance tasks included in the FUNfitness test battery (i.e., feasibility). The second aim is to investigate whether there are differences in performance on the strength and balance subtests when people with ID perform twice with a time-interval of two weeks between test sessions (i.e., test-retest reliability). This study is the first attempt to improve the quality of the data collection and to develop guidelines for optimizing future SO FUNfitness assessments.

Materials and Methods

Participants

A sample of 36 children, adolescents and young adults with mild to moderate ID (22 males and 14 females, *M* age = 16.1 years, *SD* = 4.7 years, range = 8–30 years) were recruited from a local sports club offering adapted sports activities for people with ID. The participants for this study participated in one of the three sport disciplines provided by the Centre of Adapted Sports on a weekly basis, which include: badminton (*n*=11), gymnastics (*n*=13) and soccer (*n*=12). The participants did not have any experience with the SO FUNfitness test battery prior to this study. The presence of having an ID

was confirmed by the parents or caregivers. General issues regarding the health status were checked before participation in the study. The health history included questions regarding previous experience of falls and use of medication. Informed written consent was obtained for all participants or their legal guardians prior to participation in the study. The study was approved by the Medical Ethics Committee of the KU Leuven.

Study design and instruments

This cross-sectional observational study consisted of two tests sessions with a battery of seven fitness tests, including three balance tests (i.e., the single-leg stance eyes open, single leg stance eyes closed, and the functional reach test) and four muscle strength tests (i.e., timed-stands test, partial sit-up test, seated push-up, and hand-grip test). The time-interval between test and retest was two weeks.

Prior to the start of data collection, six test administrators with background in physiotherapy and/or adapted physical activity were trained to ensure standardization in test taking, providing instructions, feedback, and scoring. During the practice sessions consensus was reached among the test administrators regarding any aspect of the manual where subjective interpretation was possible.

The actual test sessions were organised as part of the weekly training practice of the participants, in a separate testing room. Every participant was tested by the same test administrator during test and retest. All subtests were administered in accordance with the instructions presented in the user's manual provided by SO (Special Olympics, 2013). During the actual test, the test administrator provided a demonstration and at least one practice trial before every subtest. Positive encouraging feedback was provided to the participants as much as possible, except during the subtests in which maximal concentration was required to optimize performance (e.g., single leg stance). The instructions employed simple, clear, and specific language and demonstrations to facilitate comprehension for the target population. Total administration time was approximately 15–20 minutes per athlete per test session.

Test battery

The single-leg stance test with eyes open (SLS-EO) measures static balance control with the assistance of visual cues. The participant was asked to stand upright with both arms on the hips. Upon the start signal of the test instructor, the participant raises one leg (knee angle 90°) while trying to maintain balance on the dominant leg as long as possible. The test continued until either the participant lost balance (e.g., removing the arms from the hips permanently, moving the standing leg or putting a feet on the ground) or completed a full 60-s trial. Afterwards, the test was repeated with the non-dominant leg. The score of the test was the total balance time in seconds for each leg separately (max 60 s).

The single-leg stance test with eyes closed (SLS-EC) measures static balance control without assistance of visual cues, by using the same procedure as in the SLS-EO. The participant was asked to close the eyes during the test, no blindfold was used.

The functional reach test (FRT) is a measure of the ability to shift your body mass by bending forwards as far as possible without taking a step and represents the functional limits of stability. Before the start of this test, the participant was asked to stand on both legs, positioned shoulder width apart behind a line taped on the floor, and shoulders in a neutral position (no protraction or retraction), and with both arms lifted 90°, elbows and fingers extended. The test administrator reads the starting position (in cm) from a tape measure attached to the wall at shoulder height. The participant is then asked to reach the arms forward as far as possible without losing balance or without movements of the lower limbs. The distance between the starting position and the maximal reaching positon was recorded (in cm). The test was repeated with the participant standing with his opposite arm closest to the wall. For data analysis, the average score of both measurements was used.

The timed-stands test (TST) measures the functional muscle strength of the lower extremities (hip and knee extension). The participant was asked to sit on a straight-back chair with hips, knees and elbows in a 90 degrees flexion and arms held besides the body. The task was to complete 10 full stands (legs fully extended), as fast as possible, from a seating position and without using their arms.

The timer stops after the participant sits down from the tenth stand. The total time needed to complete the trial in seconds was recorded. The test was performed after demonstration and practice.

The partial sit-up test (PSUT) is a measure of the abdominal muscle strength and endurance. The participant was lying in supine position on a mat with hips and knees 90° flexed (legs placed on a chair). Their arms were held in front of the chest with elbows fully extended. The test administrator asked the participant to lift the head, sit-up to touch the chair, and then go back to the starting position. They attempted to do as many sit-ups as possible within one minute. After one minute had elapsed, the number of completed sit-ups was recorded.

The hand-grip test (HGT) is a measure of the maximum isometric muscle strength of the forearm and the hand. The participant was asked to hold the electronic dynamometer (Baseline, model 12-0286), with the elbow 90° flexed, and squeeze it as hard as possible for about 6 s, without moving the other body parts. The task was repeated three times on each side, alternating between the left and right hands. The muscle strength of the six attempts was recorded in kilograms. The best scores for the dominant hand and the non-dominant hand were used for further analysis.

The seated push-up test (SPUT) is a measure of the functional muscle strength of the triceps, shoulder and scapular muscles. The participant was positioned on the floor with lower limbs outstretched, and heels resting on the mat. While holding on to push-up blocks, the participant lifted the upper body off the floor until the elbows are straight. They held this position for as long as possible. The maximum time in seconds that they could hold the push-up position was recorded. To neutralize the effect of fatigue, the tests were performed in a logical and standardized sequence, alternating balance and strength measures: TST, SL-EO, SL-EC, PSUT, HGT, SPUT and FRT.

Statistical analysis

The statistical analysis was performed with IBM SPSS Statistics Version 24.0. We analysed the test-retest reliability by calculating the intraclass correlation coefficient (ICC) for all repeated tests. A two-way mixed model of ANOVA was used. Although there are no standard values for acceptable ICC, some guidelines are available. Values above 0.75 represent good reliability but reliability mostly should exceed 0.90 to be desirable (Portney & Watkins, 1993). We determined the standard error of measurement (SEM) to evaluate the stability of the responsiveness of each test, using Equation 1, where SD_{test1} and SD_{test2} are the standard deviations of the participants' test scores on test and retest, respectively.

$$\left(\frac{SD_{test1} \times SD_{test2}}{2}\right) \times \sqrt{1 - ICC} \quad (1)$$

The SEM% was calculated using Equation 2, where $M1$ and $M2$ are the mean values of the participants' test scores on test and retest, respectively.

$$\frac{SEM}{\left(\frac{M_{test1} + M_{test2}}{2}\right)} \times 100 \quad (2)$$

To define the amount of change that reflects a true difference between the two test sessions, we calculated the smallest real difference (SRD) by 1.96 x SEM x √2 within a 95% confidence interval and the SRD% was calculated with Equation 3.

$$\frac{SRD}{\left(\frac{M_{test1} + M_{test2}}{2}\right)} \times 100 \quad (3)$$

The *P*-value was set at .05. We compared the difference between test and retest with a paired *t*-test analysis. To compare the variance between the two tests, we provide Bland-Altman plots. A Bland-Altman plot visualizes the comparison between the test and the retest outcomes. It plots the differences between the two measures against the averages of the two measures. The mean and the highest and lowest border within the 95% confidence interval are visualized within the graph.

Results

Feasibility

A complete dataset with scores on all seven subtests on both test occasions was available for 30 ID participants. Three participants performed the retest after a three-week time interval, because they were not able to attend their training session when the retest was initially planned. Three participants dropped out after the first test, because they were not able to attend the first or second retest occasion. All participants with ID were able to understand the instructions (including demonstration and practice trial) of all the subtests. The final test (i.e., FRT) was not performed by one participant during the first test, and for two participants in the retest phase, because of lack of sustained motivation to complete the full test battery. For one participant, it was impossible to perform the 'standing on one leg' balance test with eyes open, and for seven participants, it was too difficult to perform the 'standing on one leg' balance test with eyes closed; therefore, they received a score of zero on the respective subtest (i.e., floor effect). Overall, the participants reported a positive experience participating in this study.

From the participants recruited, 30 participants with ID (83.3%) were able to attend both testing occasions. Except for SLS-EO, SLS-EC, and FR, a completion rate of 100% was noted for participants who were present for test and retest sessions. The FRT, the final test performed during testing, was completed on both test occasions by 90% of the participants. Non-completion of the FRT was due to lack of sustained motivation to complete the full test battery. The SLS-EO and SLS-EC had a completion rate of 96.7% and 76.7%, respectively, because of some participant's inability to stand on one leg.

Test-retest reliability

In Table 1, the mean values and standard deviations of both test and retest are reported, as well as ICC, SEM and SRD at a confidence interval of 95% of all tests. ICC values indicated good test-retest reliability (0.75–0.89) for all tests. The SEM values of every test were of acceptable measurement precision (SEM < SD/2). There were no significant differences from paired *t*-test between the test and retest for any of the subtests (P-values are ranging between p=0.10 for SPUT, and p=0.94 for OLS-EC).

The Bland-Altman plots for the balance and strength subtests are shown in Figures 1 and 2, respectively. Despite some outliers in every test, no major systematic bias was indicated in the plots. The dispersion around the Bland-Altman is randomly spread.

Discussion

The purpose of this study was to investigate the feasibility and the test-retest reliability of the balance and muscle strength tests used in the SO FUNfitness test battery. All of the subtests were performed in standardized conditions. The tests demonstrated good test-retest reliability, stable responsiveness, no differences between test and retest performance, and no major systematic bias. In terms of feasibility, a completion rate of 100% was observed in most subtests. We attribute this to the tests having instructions that were easy to understand for people with ID; however, the balance tests, mainly the single leg stance with eyes closed appeared to be too difficult to perform for one in every four of the participants.

With respect to the test-retest reliability of the balance tests, our findings are comparable with the findings by Blomqvist, Wester, Sundelin, and Rehn (2012) who investigated test-retest reliability of functional balance tests in a population of 89 adolescents (age range 16 to 20 years) with mild to moderate ID in a special school in Sweden. The ICC value of the SLS-EO (0.88) is identical to the values found in our study (0.87 and 0.88, respectively for the dominant side and non-dominant side). For the FRT, they used a modified version, in which the participants had to push a metal plate while reaching forward, resulting in higher ICC (0.80) compared to the ICC value found in the present study (0.75). The modification was made by Blomqvist and colleagues (2012) because they

experienced that people with ID had problems understanding the original functional reach test. Also during the administration of the FUNfitness test battery in our pilot study,   the test administrators experienced difficulties with the original test instructions (without metal plate) as written in the SO-manual; so we suggest using the device described by Blomqvist when doing the FRT to increase standardization and enhance feasibility.

Table 1. Test- retest reliability of muscle strength and balance tests in persons with ID.

	Mean (± SD)		ICC (95% CI)	SEM	SEM%	SRD	SRD%
	Test	Retest					
SL-EO dominant (s)	26.51 (23.33)	28.29 (24.56)	0.87 (0.76-0.94)	8.50	7.76	23.56	21.50
SL-EC non- dominant (s)	19.32 (20.56)	19.30 (21.11)	0.88 (0.77-0.94)	7.19	9.31	19.92	25.80
SL-EC dominant (s)	8.37 (11.34)	7.28 (11.34)	0.88 (0.76-0.94)	3.91	12.49	10.84	34.63
SL-EC non- dominant (s)	4.35 (8.55)	3.38 (4.53)	0.89 (0.78-0.94)	2.19	14.16	6.07	39.24
FRT (cm)	27.14 (11.50)	24.39 (10.18)	0.75 (0.54-0.87)	5.45	5.29	15.11	14.66
TST (s)	18.98 (8.47)	18.62 (7.23)	0.84 (0.67-0.92)	3.16	16.80	8.76	11.64
PSUT (number)	24.31 (11.65)	23.61 (14.58)	0.83 (0.68-0.91)	5.42	22.64	15.03	15.69
HGT dominant (kg)	24.25 (11.07)	26.26 (10.53)	0.86 (0.73-0.93)	4.00	15.83	11.08	10.97
HGT non- dominant (kg)	22.25 (9.85)	22.67 (8.96)	0.88 (0.77-0.94)	3.26	14.51	9.03	10.05
SPUT (s)	35.43 (22.57)	37.93 (23.75)	0.87 (0.75-0.94)	8.25	22.50	22.88	15.59

SL-EO: Single leg stance - eyes open; SL-EC: single leg stance eyes closed; FRT: Functional reach test; TST: Timed-stands test; PSUT: Partial sit- up test; HGT: Handgrip test; SPUT: Seated push up test; SEM: standard error of measurement; ICC: intraclass correlation coefficient; SRD: smallest real difference; ID: intellectual disability.
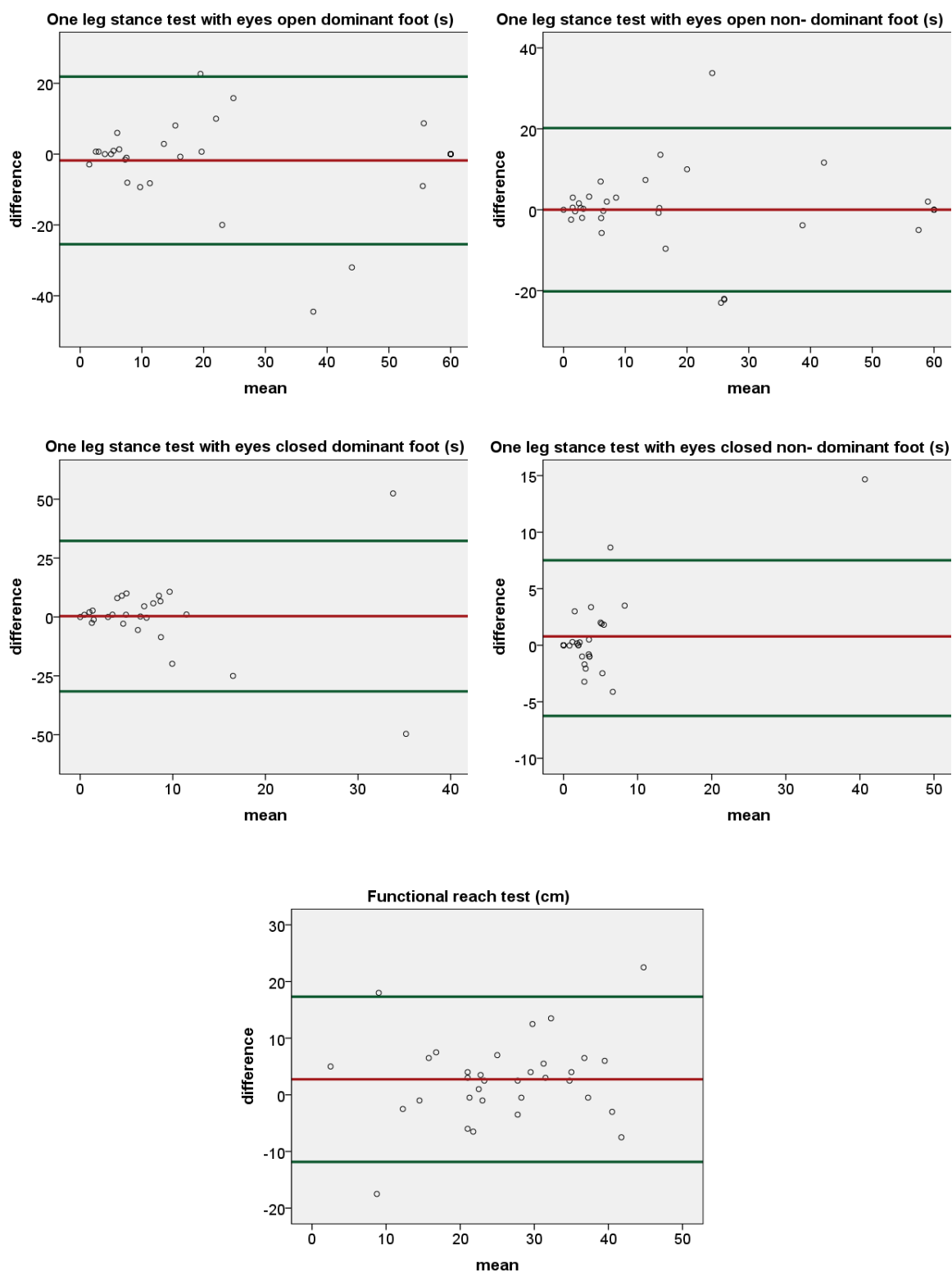
Figure 1. Bland-Altman plots for balance subtests. The center line represents the mean difference between test and retest and the outer lines are the highest and lowest border of the 95% confidence interval of the mean.
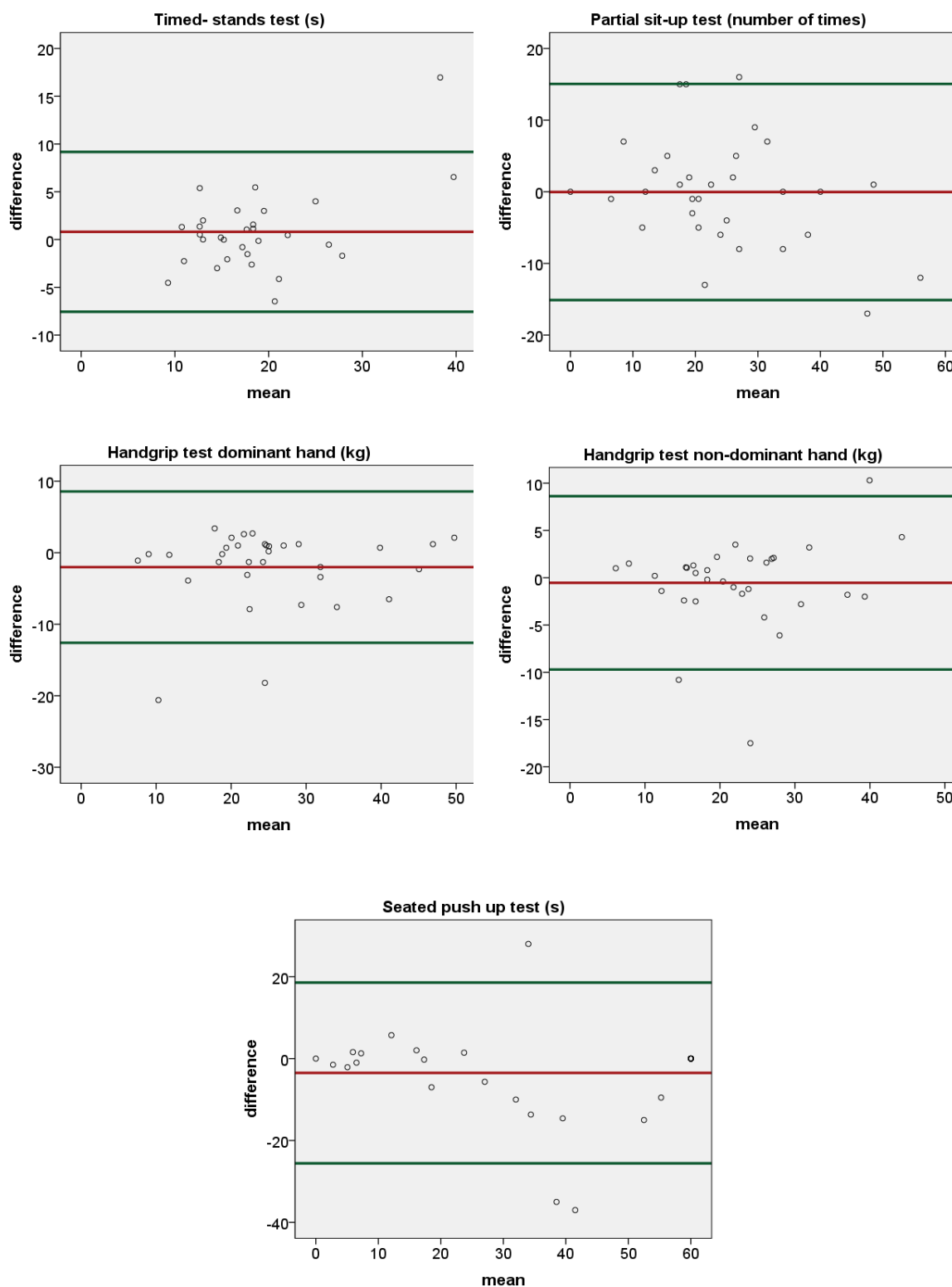
Figure 2. Bland-Altman plots for the muscle strength subtests. The center line represents the mean difference between test and retest and the outer lines are the highest and lowest border of the 95% confidence interval of the mean.

In another study, performed by Boer and Moss (2016), the test-retest reliability of 12 functional fitness tests was investigated in a sample of 43 South African adults (age range 18 to 50 years) with Down Syndrome. Balance testing in their study included a subtest comparable to the SLS-EO (ICC 0.93 and 0.98 for the right leg and left leg respectively). The major difference in test procedures for the single leg stance test was the test duration and associated maximum score. Whereas the test was completed and maximum score awarded after 60 s in our study, the test completion time was set at 10 s in the study by Boer and Moss (2016). The higher ICC values (0.93 and 0.98, respectively for the right and left leg) in their study could likely be related to the difference in test procedures, and the associated ceiling effect. Looking at the performance on the single leg stance test with the dominant leg, the average score in our study (28.9 s) was far above their maximum score of 10 s.

Aside from balance tests, Boer and Moss (2016) also included five muscle strength tests, of which four were identical or at least comparable with the TST, PSUT, HGT and SPUT used in the SO FUNfitness test battery. The ICC values they found on the muscle strength subtests varied between 0.94–0.99, which is higher compared to the ICC between 0.83–0.87 in our study. The ICC values reported in the study by Hilgenkamp and colleagues (2012) among 36 elderly people with ID (age range 50–89) were 0.90 for the HGT and 0.65 for the TST. Possible explanations for the observed differences in ICC values between studies are the characteristics of the population (age, aetiology of the impairment, cultural differences) and reported variations in test procedures, scoring, and equipment, for example the use of a different type of hand grip dynamometer, and the use of handles for the SPUT. For the TST, for example, the scoring in the FUNfitness test battery was recorded as time needed to complete 10 repetitions, whereas the chair stand test in Boer and Moss (2016) and Hilgenkamp et al. (2012) is recoding the number of repetitions in 30 s.

A limitation of the current study is that the exact level of ID of the participants was not determined whereas it would have been worthwhile to analyse it as a possible confounding factor during completion of the tests. Another limitation is the relatively small sample of participants with a broad age range (8-30 years old). Participants were recruited on the basis of eligibility criteria to compete in SO events (= minimal 8 years old). The high variance in the ages of the participants might have contributed to the higher ICC values compared to other studies that had lower variance in the ages of their participants. Furthermore, whereas standardized conditions were taken care of as much as possible, it is not always possible when working with a population of people with ID to follow the procedures without any deviations. During the testing sessions, the test administrators agreed on a certain range of flexibility in terms of testing trials, i.e., when people needed more time for practice than the one trial that was foreseen in the protocol, this was allowed. Finally, the retest took place after three weeks instead of two weeks for three of the participants because they were unable to attend the initially planned session, which might have confounded the results.

As already stated before, this study was performed in standardized conditions, whereas the intention was to analyse the test-retest reliability and feasibility of the SO FUNfitness test battery. Therefore, it is crucial to consider, when interpreting these results, that actual testing conditions during SO events are less standardized. Further research is necessary to determine whether these balance and muscle strength tests are also reliable in the specific setting of SO. Actual testing conditions observed during the previous two editions of the Belgian SO National Games in 2016 and 2017 deviated from the standardized conditions in this study. For example, the SO volunteers are not all experienced test administrators, with some volunteers only receiving very limited training and practice immediately before the start of the testing day, resulting in deviations from the standardized instructions, lack of demonstration and practice trials for the participants prior to the actual test. Furthermore, the testing environment during SO was crowded and noisy, with many distractions. As it has been demonstrated that persons with ID often have concentration problems, these factors could contribute to a reduced reliability (Hastings, Beck, Daley, & Hill, 2005; Simonoff, Pickles, Wood, Gringras, & Chadwick, 2007).

Recommendations for future data collection within the SO Healthy Athletes program are to improve the FUNfitness manual by including directions on standardization and optimized testing conditions, to foresee sufficient time for training and practice of the volunteers prior to testing, and

to organize the testing in a large enough separate space free from distractions. We also recommend the use of pictograms at every testing station, serving a twofold goal, i.e., helping the athlete to understand the task, and in the meantime helping the test administrator to memorize the test protocol. Regarding the test procedures, we also recommend the future use of the modified FRT, with the participants pushing a metal plate, because this modification makes the FRT easier to understand and to perform.

The focus in this study was on the balance and strength assessment only, whereas the complete FUNfitness test battery also includes flexibility and aerobic fitness measures. Future studies should also investigate the psychometric properties of these measures.

Conclusions

The results of this study yielded adequate test-retest reliability for the balance and muscle strength tests used as part of the FUNfitness test battery within the SO Healthy Athletes program. The testing conditions were optimized for this study to guarantee standardized test procedures.

Perspectives

There is a need for highly valid and reliable test scores to address the fitness of people with ID for many purposes. To make future field-based data collection more reliable, it is crucial to consider the recommendations enhancing standardization, and to consider the use of the modified FRT test. A recent paper by Lloyd, Foley, and Temple (2018) highlighted the uniqueness and relevance of the SO Healthy athletes' database, of which the FUNfitness test battery is an integral part. To maximize the future use of these valuable data for the purpose of research and policymaking, and to increase our knowledge and understanding of the health of individuals with ID, there is a high need for a solid evidence-base.

Authors affiliations:

[1] Faculty of Kinesiology and Rehabilitation Sciences, Department of Rehabilitation Sciences, Katholieke Universiteit Leuven, Belgium; debbie.vanbiesen@kuleuven.be

[2] Faculty of Kinesiology and Rehabilitation Sciences, Department of Rehabilitation Sciences, Katholieke Universiteit Leuven, Belgium; roicharles.pineda@kuleuven.be

* Correspondence: Debbie.vanbiesen@kuleuven.be; Tel.: +32 16 3765 12

Author Contributions: Conceptualization, D.V. and R.P.; Methodology, D.V.; Formal Analysis, D.V and R.P..; Investigation, D.V and R.P..; Writing-Original Draft Preparation, D.V.; Writing-Review & Editing, R.P.; Visualization, D.V.; Supervision, D.V.; Project Administration, D.V.

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

References

AAIDD. (2010). *Intellectual disability: Definition, classification, and systems of supports* (11th ed.). Washington, DC: American Association on Intellectual and Developmental Disabilities.

Bandini, L.G., Curtin, C., Hamad, C., Tybor, D.J., & Must, A. (2005). Prevalence of overweight in children with developmental disorders in the continuous national health and nutrition examination survey (NHANES) 1999 –2002. *Journal of Pediatrics*, *146*, 738–743. doi:10.1016/j.jpeds.2005.01.049

Blomqvist, S., Olsson, J., Wallin, L., Wester, A., & Rehn, B. (2013). Adolescents with intellectual disability have reduced postural balance and muscle performance in trunk and lower limbs compared to peers without intellectual disability. *Research in Developmental Disabilities*, *34*, 198–206. doi:10.1016/j.ridd.2012.07.008

Blomqvist, S., Wester, A., Sundelin, G., & Rehn, B. (2012). Test-retest reliability, smallest real difference and concurrent validity of six different balance tests on young people with mild to moderate intellectual disability. *Physiotherapy*, *98*, 313–319. doi:10.1016/j.physio.2011.05.006

Boer, P. H., & Moss, S. J. (2016). Test-retest reliability and minimal detectable change scores of twelve functional fitness tests in adults with Down syndrome. *Resrach in Developmental Disabilities*, *48*, 176–185. doi:10.1016/j.ridd.2015.10.022

Borji, R., Zghal, F., Zarrouk, N., Sahli, S., & Rebai, H. (2014). Individuals with intellectual disability have lower voluntary muscle activation level. *Research in Developmental Disabilities*, *35*, 3574–3581. doi:10.1016/j.ridd.2014.08.038

de Rezende, L.F., Rodrigues Lopes, M., Rey-López, J.P., Rodrigues Matsudo, V.K., & do Carmo Luiz, O. (2014). Sedentary behavior and health outcomes: an overview of systematic reviews. *PLoS One*, *9*(8), e105620. doi:10.1371/journal.pone.0105620

Enkelaar, L., Smulders, E., Van Schrojenstein Lantman - de Valk, H., Geurts, A.C., & Weerdesteyn, V. (2012). A review of balance and gait capacities in relation to falls in persons with intellectual disability. *Research in Developmental Disabilities*, *33*, 291–306. doi:10.1016/j.ridd.2011.08.028

Fjørtoft, I, Pedersen, A.V., Sigmundsson, H., & Vereijken, B. (2011). Measuring physical fitness in children who are 5 to 12 years old with a test battery that is functional and easy to administer. *Physical Therapy*, *91*, 1087–1095. doi:10.2522/ptj.20090350

Foley, J.T., Lloyd, M., Vogl, D., & Temple, V.A. (2014). Obesity trends of 8–18 year old Special Olympians: 2005-2010. *Research in Developmental Disabilities*, *35*, 705–710. doi:10.1016/j.ridd.2013.12.005

Hale, L., Bray, A., & Littmann, A. (2007). Assessing the balance capabilities of people with profound intellectual disabilities who have experienced a fall. *Journal of Intellectual Disability Research*, *51*, 260–268. doi:10.1111/j.1365-2788.2006.00873.x

Hastings, R.P., Beck, A., Daley, D., & Hill, C. (2005). Symptoms of ADHD and their correlates in children with intellectual disabilities. *Research in Developmental Disabilities*, *26*, 456–468. doi:10.1016/j.ridd.2004.10.003

Hilgenkamp, T.I., van Wijck, R., & Evenhuis, H.M. (2012). Feasibility and reliability of physical fitness tests in older adults with intellectual disability: A pilot study. *Journal of Intellectual and Developmental Disabilities*, 37, 158–162. doi:10.3109/13668250.2012.681773

Hsieh, K., Heller, T., & Miller, A.B. (2001). Risk factors for injuries and falls among adults with developmental disabilities. *Journal of Intellectual Disability Research*, *45*, 76–82. doi:10.1111/j.1365-2788.2001.00277.x

Hsieh, K., Rimmer, J., & Heller, T. (2012). Prevalence of falls and risk factors in adults with intellectual disability. *American Journal of Intellectual and Developmental Disabilities*, *117*, 442-454. doi:10.1352/1944-7558-117.6.442

Hsu, T.-Y. (2016). Effects of Wii Fit® balance game training on the balance ability of students with intellectual disabilities. *Journal of Physical Therapy Science*, *28*, 1422–1426. doi:10.1589/jpts.28.1422

Hutzler, Y., & Korsensky, O. (2010). Motivational correlates of physical activity in persons with an intellectual disability: A systematic literature review. *Journal of Intellectual Disability Research*, *54*, 767–786. doi:10.1111/j.1365-2788.2010.01313.x

Lee, K., Lee, M., & Song, C. (2016). Balance training improves postural balance, gait, and functional strength in adolescents with intellectual disabilities: Single-blinded, randomized clinical trial. *Disability and Health Journal*, *9*, 416–422. doi:10.1016/j.dhjo.2016.01.010

Lin, J.-D., Yen, C.-F., Li, C.-W., & Wu, J.-L. (2005). Patterns of obesity among children and adolescents with intellectual disabilities in Taiwan. *Journal of Applied Research in Intellectual Disabilities*, 18, 123–129. doi:10.1111/j.1468-3148.2005.00241.x

Lloyd, M., Foley, J.T., & Temple, V.A. (2018). Maximizing the use of Special Olympics International's Healthy Athletes database: A call to action. *Research in Developmental Disabilities*, *73*, 58–66. doi:10.1016/j.ridd.2017.12.009

Lotan, M., Isakov, E., Kessel, S., & Merrick, J. (2004). Physical fitness and functional ability of children with intellectual disability: Effects of a short-term daily treadmill intervention. *Scientific World Journal*, *4*, 449–457.

Portney, L., & Watkins, M. (1993). *Foundation of clinical research: Application to practice*. Philadelphia, PA: F.A. Davis.

Schützwohl, M., Koch, A., Koslowski, N., Puschner, B., Voss, E., Salize, H. J., … & Vogel, A. (2016). Mental illness, problem behaviour, needs and service use in adults with intellectual disability. *Social Psychiatry and Psychiatric Epidemiology*, *51*, 767–776. doi:10.1007/s00127-016-1197-4

Sherrard, J., Tonge, B.J., & Ozanne-Smith, J. (2001). Injury in young people with intellectual disability: Descriptive epidemiology. *Injury Prevention*, *7*, 56–61. doi:10.1136/ip.7.1.56

Simonoff, E., Pickles, A., Wood, N., Gringras, P., & Chadwick, O. (2007). ADHD symptoms in children with mild intellectual disability. *Journal of the American Academy of Child & Adolescent Psychiatry*, *46*, 591–600. doi:10.1097/chi.0b013e3180323330

Special Olympics. (2013). Special Olympics FUNfitness: Learn how to organize, promote and present. Retrieved 23 April 2017, from https://media.specialolympics.org/resources/health/disciplines/funfitness/FUNfitness-Training-Manual.pdf?_ga=2.201019054.1625048988.1524495889-426579843.1524495889

Taylor, R. W., Jones, I. E., Williams, S. M., & Goulding, A. (2002). Body fat percentages measured by dual-energy X-ray absorptiometry corresponding to recently recommended body mass index cutoffs for overweight and obesity in children and adolescents aged 3–18 y. *American Journal of Clinical Nutrition*, *76*, 1416–1421. doi:10.1093/ajcn/76.6.1416

Winnick J.P., & Short, F.X. (1999). *The Brockport Physical Fitness Test*. Champaign, IL: Human Kinetics.

Wuang, Y. P., Chang, J. J., Wang, M. H., & Lin, H. C. (2013). Test-retest reliabilities of hand-held dynamometer for lower-limb muscle strength in intellectual disabilities. *Research in Developmental Disabilities*, *34*, 2281–2290. doi:10.1016/j.ridd.2013.04.010