Harding, J, Tallent, J, Sheriff, K, McCann, C, Cortes, N, Olsson, L, Shaw, J and Howe, L

The Reliability of Physical Performance Testing within Elite Adolescent Pre-Professional Ballet Dancers

The Reliability of Physical Performance Testing within Elite Adolescent Pre-Professional Ballet Dancers

Jamie Harding[1,2], Jamie Tallent[1,3], Karen Sheriff[2], Chris McCann[2], Nelson Cortes[1,4], Luke Olsson[5], Joseph Shaw[6,7], **Louis Howe[1]**

[1]School of Sport, Exercise and Rehabilitation Sciences, University of Essex, Essex, UK.

[2]Royal Ballet School Healthcare Team, Royal Ballet School, London, UK.

[3]Department of Physiotherapy, Faculty of Medicine, Nursing and Health Science, School of Primary and Allied Health Care, Monash University, Melbourne, VIC, Australia.

[4]Deparment of Bioengineering, George Mason University, Fairfax VA UK

[5]School of Sport and Exercise Sciences, Liverpool John Moores University, Liverpool, UK.

[6]Royal Ballet Healthcare, The Royal Ballet, The Royal Opera House, London, UK

[7]Faculty of Sport, Technology and Health Sciences, St Mary's University, Twickenham, London, UK

Address for correspondence:

Dr Louis Howe

University of Essex, Wivenhoe Campus, Colchester, Essex, United Kingdom.

CO43SQ

Tel: +441206 873333

Email: louis.howe@essex.ac.uk

Acknowledgments: none

Declaration of conflicting interest: none

Ethical approval and informed consent: Written consent was obtained from parents for all participants for data collection and publication, while ethical approval was provided by the University of Essex Ethics Committee.

Funding statement: This project was joint funded by the University of Essex and Royal Ballet School.

The Reliability of Physical Performance Testing within Elite Adolescent Pre-Professional Ballet Dancers

Abstract

Introduction: Evaluating and training strength qualities is crucial for the physical development of ballet dancers. Whilst data is available as to the sensitivity of strength tests for detecting changes in athlete populations, between-session reliability for adolescent ballet dancers is yet to be determined. This study aimed to determine the between-session reliability of physical performance tests in elite adolescent ballet dancers. **Methods:** Depending on the test, a cohort of 25 to 54 pre-professional ballet dancers (9 to 30 males, 14 to 29 females) participated in a series of six physical tests across 12 sessions. Each testing session involved performing one strength test, with retesting administered seven days later. The testing protocol included single-leg isometric squat, single-leg isometric plantarflexion, countermovement jump, standing single-leg countermovement jump, drop jump from 30 cm, and for males, seated overhead press to voluntary failure using 30 kg. Data was analysed using a pairs sample t-test, interclass correlation coefficients and measures of absolute reliability including values of minimal detectable change. **Results:** Pairs sample t-tests revealed no systematic bias was present between trial 1 and 2 for each test**.** Across all tests, interclass correlation coefficients ranged from *good* to *excellent* (0.89-0.98), and coefficients of variation were 2.6–6.5%. **Conclusion:** These results indicate strength testing can reliably be integrated into a comprehensive physical performance testing battery to identify changes associated with improved physical performance across the academic year for adolescent ballet dancers. Based on the minimum detectable change values, changes in jump performance across the range of tests employed in this study can likely be detected after relatively short training periods. However, maximal isometric strength tests such as the single-leg squat may require longer than six weeks to detect performance changes. The current study expands the testing options for ballet training centres and high-performance settings, ensuring confidence in accurately measuring physical changes.

Key words: ballet, adolescent dancers, strength testing.

Introduction

The physical demands of elite ballet are considerable, with hours spent dedicated to classes, rehearsals and performances surpassing those observed in athletic populations.[1] Consequently, well-developed strength qualities are required to allow ballet dancers to maintain proper technical alignment, balance, and stability throughout a range of balletic movements, providing the foundation to height in jumps, extension in leg lifts, efficiency in overhead lifts, and stability in turns.[2] Furthermore, dancers’ strength qualities underpin the ability to rapidly produce force, facilitating quick transitions between steps, accelerations, and leaps during allegro sequences performed in the studio and on stage.[3] Common methods used in various high performance athletic populations for evaluating lower and upper body strength qualities include testing maximal isometric force production,[4] muscular endurance[5] and jump performance.[6] The utilisation of such tests in ballet may offer valuable insights into dancers' physical capabilities and facilitate the design of tailored training programs, performance tracking, and injury risk management.[7]

The reliability of performance tests in athletic populations have been documented.[8,9] For example, Carroll et al identified an intraclass correlation coefficient (ICC) of 0.92 and a coefficient of variation (CV%) of 3.2% when examining performance during the countermovement jump test in Division-I college athletes.[6] Similarly, Blagrove et al and McGoldrick et al reported *good* to *excellent* reliability with high sensitivity for maximal isometric strength testing (ICC = 0.86-0.92, CV% = 4.4-8.4%) in adolescent distance runners and youth soccer players.[10,4] While this data may inform the interpretation for performance testing in youth athletes, ballet dancers exhibit distinct motor skills and physiological adaptations owing to the aesthetic nature of ballet, making ballet clear and distinct from more objective based high-performance activities.[11,12,13] This may result in divergent performance outcomes in strength tests due to set coordination patterns observed during activities like jump-landings. To date, reliability studies for strength testing in ballet have primarily focused on adult populations, with only 42% of participants in the Mattiussi et al study being elite ballet dancers, while the remaining participants were active individuals.[14] Kolokythas et al. tested elite adolescent ballet dancers but only evaluated one isometric strength test.[15] Consequently, the error associated with a range of physical performance tests for adolescent ballet dancers is unknown and needs further investigation. Reliability data derived from a ballet population will better inform practitioners supporting the physical development of ballet dancers by helping to distinguish between potential 'noise' and actual changes in test performance.

Bilateral strength tests involving both legs have been the traditional approach for evaluating lower body strength in sports medicine.[16] However, determining performance using single-leg strength assessments may provide novel insights into force production capabilities.[17] For example, unilateral tests may offer further valuable insights into limb strength characteristics, particularly useful when establishing criteria for return-to-dance protocols following unilateral injuries or directing training emphasis in non-injured individuals with potential performance asymmetries.[18] Additionally, unilateral maximal isometric strength tests may provide a more accurate representation of an individual's maximal strength when compared to bilateral testing during standing tests, as the tolerance for spinal loading may no longer be the limiting factor for global force output.[14] Due to the scarcity of research employing unilateral strength testing among elite adolescent ballet dancers this necessitates further investigation to enhance practical insights.[15,19] When analysing the jumping demands of classical ballet, research has only recently quantified the loading associated with ballet training, highlighting that junior dancers perform a higher number of jumps than senior dancers, and males jump at a greater volume than females.[20] With jump counts during class ranging from 62 to 270—exceeding those reported in other jumping-based sports such as basketball and volleyball[20]—monitoring jumping performance is crucial not only for optimising performance but also for injury management. Given that jumping tasks account for over 50% of injury-related time loss in ballet companies, tracking jump performance can serve as both a performance metric and a key marker for return to full balletic training following injury.[20] Furthermore, from an artistic perspective, ballet company directors, choreographers, senior teachers, and experienced dancers regard power and jumping ability as essential attributes for success in professional ballet, underscoring the need for objective monitoring of jumping performance.[3]

There is a need to consider gender-specific physical tests in ballet, given the differences in movement demands. For example, male dancers engage in extensive overhead lifting during performance and training[2,21] and, therefore, measures of upper limb strength are important to inform programming for this population.[5] Additionally, male ballet dancers have an elevated risk of lower back injuries as a consequence of repetitively performing a high volume of lifts, necessitating an objective measure of upper body muscular endurance to inform attempts to mitigate the prevalent injury risk of the lower back, as highlighted by artistic and healthcare professionals in ballet settings.[3] As evidence for the accuracy of testing overhead lifting strength in male ballet dancers is currently limited, there is a need to investigate the reliability of testing overhead lifting capability.[22]

Muscular strength is crucial to performance in ballet[2], suggested as a critical trait for a ballet dancer to possess by artistic staff when selecting prospective ballet dancers.[3] Therefore, determining the sensitivity of testing protocols will support practitioners in designing impactful training programs for dancers. However, currently there is a lack of evidence concerning their use in high-level dance environments, especially among adolescent dancers. Therefore, this study aims to determine the measurement error between sessions when testing single-leg isometric squat, standing single-leg isometric plantarflexion, countermovement jump, single-leg countermovement jump, drop jump, and for males, seated overhead press to voluntary failure adolescent elite ballet dancers.

Methods

Study Design

A between-session repeated measures design was used to determine the inter-session reliability of performance tests in pre-professional ballet dancers. Dancers reported to the Strength and Conditioning facility at the Royal Ballet School, with one test performed in each testing session. Re-testing was performed seven days later at the same time of day, before classes had started, to account for variations in circadian rhythm[23] and timetable demands. With six strength tests included in the physical performance testing battery, testing occurred over a 12-week period (Figure 1). Performance tests included the single leg isometric squat (SL squat), standing single leg isometric plantarflexion (SL PF), seated overhead press repetitions to volitional fatigue with 30kg (OHP), countermovement jump (CMJ), single leg countermovement jump (SL CMJ), drop jump (DJ) tests. Prior to each testing session, a standardised warm up was performed.

*****Insert Figure 1*****

Participants

A priori power analyses were performed using the calculation outlined by Walter et al. (1998)[24] indicating that a minimum of twenty-three participants were required to detect the minimal acceptable reliability of ICC values of 0.7. This calculation was based on a significance level (α) of 0.05 and a power (β) of 80%, aiming to reach the expected reliability of ICC values greater than 0.9.[6,14] Due to the 12-week data collection period, not all participants completed both sessions for each test. Consequently, the participant characteristics vary for each test and are summarized in Table 1.

All participants were screened prior to testing to ensure physical health, with injured participants or recently injured participants (an injury was defined as a musculoskeletal condition that hindered normal training activities within the week leading up to data collection) excluded from data collection. Written consent was obtained from parents for all participants and ethical approval provided by the University of Essex Ethics Committee.

Table 1. Participant information for each test.

Test	*n*	*n* (male)	*n* (female)	Age (years)	Maturity Offset (years)	Height (cm)	Mass (kg)
Drop jump	44	25	19	17 ± 1	3.0 ± 1.3	171.8 ± 8.7	57.8 ± 8.8
Countermovement jump	59	30	29	17 ± 2	3.0 ± 1.8	174.4 ± 8.4	57.1 ± 8.9
Single-leg countermovement jump	54	28	26	17 ± 2	2.9 ± 1.8	171.4 ± 8.2	57.1 ± 8.5
Single-leg isometric squat	25	9	16	17 ± 1	3.5 ± 0.8	169.5 ± 6.1	54.8 ± 7.0
Single-leg isometric plantarflexion	26	12	14	17 ± 1	3.4 ± 0.9	171.6 ± 8.8	57.6 ± 8.6
Seated overhead press with 30kg	25	25	0	17 ± 1	2.4 ± 1.5	178.2 ± 6.0	65.4 ± 7.2

Procedures

All participants were familiarised with the physical performance test before data collection having performed the tests in previous physical profiling sessions and given the option of a practice attempt before any data was collected. Coaching was provided where appropriate to ensure technical proficiency, data collection was initiated once dancers had verbalised they understood the protocol and were confident in performing the test. Testing sessions began with a 5-minute standardised, progressive warm-up involving 2 sets of 10 repetitions of bodyweight squat, hip hinge, calf raise, banded vertical pressing movements as well as 2 sets of 10 repetitions of pogo jumps, 2 sets of 5 repetitions of single leg countermovement jumps and 2 sets of 5 repetitions of countermovement jump requiring submaximal efforts. All unilateral tests were collected on the left limb first, followed by the right limb to standardise the order of contractions. All isometric and jump tests were conducted barefoot on a force platform (ForceDecks 4000, VALD Performance, Queensland, Australia) sampling at 1000 Hz. For all isometric strength tests, a custom isometric rig with 2.5 cm adjustable vertical spacing and a barbell (Original 2028 Olympic Bar, Strength Shop, United Kingdom) were used, with a 5 cm thick foam pad (Olympic Neck Pad, Perform Better, United Kingdom) placed around the barbell for participant comfort. The vertical ground reaction force data acquired from each jump and the maximal isometric strength tests were analysed via the ForceDecks software (ForceDecks, VALD Performance, Queensland, Australia). Prior to the initial testing session for each test, bodyweight was collected during a static trial during which participants stood motionless on the force platform. Standing and seated height were collected one week before data collection using a medical grade measuring station (Seca 287 Wireless Ultrasonic Measuring Station, Hamburg, Germany). Maturity offset calculations were estimated using non-invasive anthropometric measures recording of each participant's gender, date of birth, standing stature, seated statue, and bodyweight.[25] Maturity offset can be defined as the as the time before or after PHV.[26] Data was collected by a trained nurse with extensive experience of collecting anthropometric data in adolescent populations.

Drop Jump

Utilising two force platforms, participants completed three DJs with approximately a 1-minute rest interval between each trial. Participants stood on a 30 cm platform with their feet hip-width apart and hands placed on the hips. To initiate the DJ, participants stepped forward from the box before landing with both feet simultaneously on the force platforms. Upon landing, participants executed a maximal rebound vertical jump while maintaining hand contact with hips throughout. Participants were cued to "jump as high and as quickly as you can, spending as little time on the floor as possible by imagining the floor is hot like lava" before performing each test.[27] Participants had the option of a practice jump before data collection, followed by an additional 1-minute rest period. The recorded metrics included jump height in centimetres, calculated via the flight-time method (calculated via the ForceDecks software), ground contact time (the duration spent in contact with the ground between initial landing and take-off), and Reactive Strength Index (RSI), calculated using the equation of flight time divided by ground contact time. For data analysis, the mean value of the three attempts used.

Countermovement Jump

Participants performed three CMJs whilst standing on two force platforms with approximately 1-minute of rest between each attempt. Participants were instructed to stand on the force plate with their feet positioned between hip and shoulder width apart and their hands placed on their hips throughout the test. All attempts were performed to a self-selected depth and the participant was cued to "*shoot up like a rocket and jump as high as you can*" before each test.[27] Participants had the option of a practice jump before data collection, followed by an additional 1-minute rest period. Jump height was determined using the flight-time method with ForceDecks software (v2.0.7418, Vald Performance) and recorded in centimetres. The highest jump and the mean value of the three attempts used for data analysis. The flight-time method for calculating jump height was selected for its applicability in dance school environments, where basic equipment, limited budgets, and restricted access to advanced training tools are common.

Single Leg Countermovement Jump

Participants completed the SL CMJ on a single force platform, conducting three consecutive attempts with approximately 1-minute rest intervals between each attempt. Participants were instructed to descend to a depth of their choosing and were cued as above. To prevent additional leg swing from the non-jumping leg, its hip and knee were held at 90° flexion. Participants had the option of a practice jump on each leg before data collection, followed by an additional 1-minute rest period. Jump height was determined using the flight-time method with ForceDecks software and recorded in centimetres, with the highest and mean value of the three attempts used for data analysis.

Single Leg Isometric Squat

Participants stood in a partial squat position with a foam pad between their neck and the bar to ensure comfort and facilitate maximal force production, with the bar positioned to rest across the superior border of the scapular. The test foot was placed in the centre of a force platform with the hands gripping the bar using an overhand claw grip. A custom-built rig was employed to set the barbell at a height that permitted flexion of the knee and hip joints to 140°, where full extension for both the knee and hip was 180°.[14] Knee angle was determined by aligning the fulcrum of the goniometer over the lateral epicondyle of the femur, while the stable arm was positioned in line with the lateral malleolus and the mobile arm aligned with the greater trochanter. For the hip angle, the fulcrum of the goniometer was placed over the greater trochanter, with the stable arm aligned with the femur and the mobile arm aligned with the glenohumeral joint. The contralateral limb was held in 90° of hip flexion to maintain a neutral hip positioning throughout the test. Participants were instructed "*you have 5 seconds to push maximally into the barbell as hard as you can, trying to bend the barbell*" before each trial. Each trial was initiated by the researcher instructing the participants to adopt the relevant position and then counting down "*3, 2, 1, Push*", with trials lasting 5s in total. Participants performed three consecutive trials on each limb and were given approximately 10s rest between trials to reset prior to the next trial. While the optimal recovery duration between maximal isometric contractions remains debated[29], we selected a relatively brief recovery period based on both established reliability from similar protocols[14] and time constraints of testing a large cohort.

Single Leg Isometric Plantarflexion

The SL PF test was selected to represent the strength qualities of all plantar flexors[30], which are associated with jump performance.[31] Participants stood in the centre of the force platform with a foam pad between the neck and barbell positioned across the superior border of the scapular. The barbell was fixed inside a custom-built rig, with the barbell height set to account for individual variance in height. The ankle joint of the test foot was positioned at 130° of plantarflexion, measured using a goniometer with the fulcrum aligned to the lateral malleolus, the stable arm in line with the head of the fibula and the mobile arm in line with the base of the 5th metatarsal. Participants were cued to have a 'soft knee' on the test limb to prevent hyperextension at the knee joint and maintain a knee and hip flexion angle between 170° and 180°.[14] The knee angle was determined by aligning the fulcrum of the goniometer over the lateral epicondyle of the femur, with the stabilisation arm positioned in line with the lateral malleolus and the mobile arm aligned with the greater trochanter. Hip position was measured by placing the goniometer's fulcrum over the greater trochanter, aligning the stabilisation arm with the lateral epicondyle of the femur and the mobile arm with the glenohumeral joint. The contralateral limb was held at 90° of hip flexion to maintain a neutral hip positioning throughout test. Participants were instructed "*you have 5 seconds to push maximally into the barbell as hard as you can, trying to bend the barbell*" before each trial. Each trial was initiated by the researcher instructing the participants to adopt the relevant position, bracing, and then counting down "*3, 2, 1, Push*". Trials lasted 5s in total. Participants performed three consecutive trials on each limb and were given approximately 10s rest between trials to reset before the next trial.

Seated Overhead Press

A 30kg Olympic barbell, measuring 10cm in circumference and 220cm in length, was securely positioned within a squat rack, placed in front of a conventional flat weightlifting bench with a height of 40cm. The participants assumed a sitting position on the bench with their feet flat on the floor and with an upright spinal posture. Participants were then instructed to execute the OHP with their hands positioned at shoulder-width apart in the front rack position, utilising an overhand claw grip. Participants were instructed to start each repetition with the barbell positioned just above the clavicles, then press it above the crown of the head while fully extending the elbows, before returning the barbell to below the chin to complete one full repetition. To warm-up, participants completed ten repetitions with a 20kg barbell followed by a 90s rest. For testing, participants pressed the barbell overhead, safely completing as many repetitions as possible with the loaded 30kg barbell. Throughout testing, an experienced safety spotter was present behind the participant to help and assist participants if they failed the test, or the barbell path deviated significantly backwards, putting the participant at risk, with no intervention before failure. Safety spotter arms were adjusted within the squat rack just below the bottom position of the OHP for each participant to ensure if test was failed, barbell would be safely collected within the squat rack. A second tester was present to perform a double count to confirm the final number of repetitions. The tester provided verbal feedback if the barbell did not reach the required depth below the chin or fully extend the elbows, allowing participants to self-correct their form; any repetitions failing to meet the criteria were discarded from the final test results. The test was stopped by the tester when the participant was unable to maintain correct technique with cueing or when no more repetitions could be completed. Lifting cadence was self-selected, with participants instructed that the barbell had to remain in constant motion throughout the test duration. The OHP test was performed once, with the total number of successful repetitions completed used for data collection.

Statistical Analysis

For isometric strength testing, the mean vertical ground reaction force (vGRF) was extracted during static bodyweight trials. Peak vGRF was extracted during maximal isometric strength trials directly from the force platform software, with no filtering applied to vGRF data as per testing guidelines.[32] Measures of relative force being calculated as peak vGRF in Newtons being divided by body mass in kilograms. Descriptive statistics (mean ± standard deviation) were calculated for all outcome variables associated with each test. For unilateral tests, variables were calculated for both limbs. The assumption of normality was confirmed using the Shapiro-Wilk test (α = <0.05). Initially, a paired samples t-test was used to calculate systematic bias between test 1 and test 2 from each performance test.[33] Relative reliability was assessed through the calculation of CV% ((SD_pooled / X̄_1,2) × 100)[33] and using two-way mixed effects models for average measures of absolute agreement (ICC (2,*k*)) across outcome measures.[34] ICCs were reported with 95% confidence intervals and were interpreted as follows: <0.5 *poor*, 0.5–0.75 *moderate*, 0.75–0.9 *good*, and > 0.9 *excellent*.[34] Absolute reliability was calculated using SEM (SD√1-ICC)[33] and MDC (SEM*1.96*√2).[35] Statistical tests were performed using JASP statistical software package (v0.17.1, University of Amsterdam, Netherlands).

Results

There was no systematic bias found between test 1 and 2 for any variable ($p \geq 0.05$). Relative and absolute values of reliability for all measures are presented in Table 2. Relative reliability was *excellent* (ICC ≥ 0.90) for all variables except relative measures of SL squat strength on the left leg (ICC = 0.87) and RSI scores derived from the DJ test (ICC = 0.89), which demonstrated *good* relative reliability. Measures of absolute reliability are reported in Table 2 for each test measure, with CV% ranging from 2.6–5.9% for all variables.

Table 2. Between-session reliability for all performance tests in elite adolescent pre-professional ballet dancers.

Test	Outcome measure	Test 1 Mean ± SD	Test 2 Mean ± SD	Change in mean	Between test p-values	ICC (95% CI)	CV%	SEM	MDC
Drop jump	RSI ($s \cdot s^{-1}$)	1.8 ± 0.5	1.7 ± 0.4	0.1	0.110	0.89 (0.80-0.94)	5.9	0.14	0.38
	Ground contact time (s)	0.29 ± 0.07	0.30 ± 0.07	0.01	0.133	0.92 (0.87-0.95)	5.0	0.02	0.05
	Jump height (cm)	30.7 ± 5.9	29.5 ± 5.6	1.2	0.181	0.93 (0.87-0.96)	3.4	1.5	4.1
Countermovement jump	Peak jump height (cm)	31.0 ± 7.6	31.4 ± 7.6	0.4	0.870	0.97 (0.95-0.98)	3.1	1.2	3.3
	Mean jump height (cm)	30.1 ± 7.3	30.3 ± 7.2	0.2	0.993	0.96 (0.94-0.98)	3.0	1.2	3.4
Single leg countermovement jump (right)	Peak jump height (cm)	14.3 ± 4.0	14.0 ± 4.1	0.3	0.244	0.95 (0.92-0.97)	5.0	0.9	2.5
	Mean jump height (cm)	13.4 ± 3.8	13.3 ± 3.9	0.1	0.670	0.96 (0.94-0.98)	4.7	0.7	2.1
Single leg countermovement jump (left)	Peak jump height (cm)	14.8 ± 4.3	14.6 ± 4.2	0.2	0.578	0.96 (0.93-0.97)	4.1	0.8	2.3
	Mean jump height (cm)	13.8 ± 3.9	13.7 ± 4.1	0.1	0.655	0.98 (0.96-0.99)	3.8	0.6	1.7
Single-leg isometric squat (right)	Absolute vGRF (N)	1663.8 ± 403.1	1665.0 ± 417.0	1.2	0.392	0.93 (0.88-0.96)	4.4	103	285
	Relative vGRF ($N \cdot kg^{-1}$)	29.7 ± 5.8	29.7 ± 5.4	0.0	0.343	0.90 (0.82-0.94)	4.5	1.8	5.0
Single-leg isometric squat (left)	Absolute vGRF (N)	1604.9 ± 370.5	1569.9 ± 391.4	35	0.980	0.91 (0.86-0.95)	4.8	119	330
	Relative vGRF ($N \cdot kg^{-1}$)	28.0 ± 5.9	28.7 ± 5.3	0.7	0.990	0.87 (0.79-0.93)	4.9	2.0	5.5
Single-leg isometric plantarflexion (right)	Absolute vGRF (N)	1561.1 ± 340.2	1560.1 ± 415.7	0.1	0.966	0.97 (0.96-0.98)	2.9	61	168
	Relative vGRF ($N \cdot kg^{-1}$)	26.3 ± 3.9	26.3 ± 5.1	0.0	0.950	0.96 (0.93-0.97)	3.0	1.0	2.7
Single-leg isometric plantarflexion (left)	Absolute vGRF (N)	1601.3 ± 372.0	1616.7 ± 379.2	15.4	0.500	0.98 (0.96-0.99)	2.6	56	156
	Relative vGRF ($N \cdot kg^{-1}$)	26.9 ± 4.0	27.2 ± 4.5	0.3	0.411	0.94 (0.90-0.97)	2.7	1.0	2.8
Seated overhead press (30kg)	Number of repetitions performed	19 ± 8	20 ± 8	1	0.331	0.98 (0.96-0.99)	6.5	1	3

vGRF = Vertical ground reaction force; ICC = Intraclass correlation coefficient; SEM = Standard error of measurement; MDC = Minimal detectable change

Discussion

This study aimed to establish the between-session reliability for a testing battery examining physical performance in elite adolescent ballet dancers. The results show that measures representing performance during lower extremity maximal isometric force production, jumping and upper extremity strength endurance tests demonstrate *good* to *excellent* relative reliability and CV% ranging from 2.6-5.9%. Hence, strength tests can be reliably incorporated into a comprehensive performance testing battery to detect performance changes typically associated with strength gains observed following a training intervention in this population. Within the between-session design, no systematic bias was observed between tests, indicating the absence of learning effects, participant bias, or acute adaptations.[33] These results imply that the procedures employed in this study are suitable for minimising the effects of systematic error.

This investigation assessed the reliability of the DJ test, *good* reliability was observed for RSI (ICC = 0.89, CV% = 5.9%), while ground contact time (ICC = 0.92, CV% = 5.0%), and jump height (ICC = 0.93, CV% = 3.4%) demonstrated *excellent* reliability. When contrasted with other studies exploring the reliability of DJ performance from a 30 cm drop height, Xu et al. (2023) reported comparable findings, with *excellent* between-session reliability for jump height (ICC = 0.95, CV = 5.4%), ground contact time (ICC = 0.97, CV = 5.9%), and RSI (ICC = 0.95, CV = 7.7%).[36] This result was unexpected, as we anticipated greater variation in drop jump performance among dancers. This expectation was based on the unique landing strategies dancers employ in ballet to meet artistic demands, particularly the pronounced ankle plantarflexion used during initial ground contact[37], potentially affecting force production relying on a fast stretch-shortening cycle.[38] Additionally, as the DJ test is not a widely used test within ballet, the novel exposure to this task combined with a unique landing strategy may increase between-session variance in jump performance.[39] This may be further evident if collecting data via equipment utilising optical sensor technology when comparing to force plate data, as landing and take-off technique may affect comparisons in jump height.[40] However, the results of this study indicate that practitioners working with dancers should expect similar variance in drop jump test performance as seen in other populations. From a practical perspective, MDC values from this study appear sensitive enough to identify performance improvements after a 12-week plyometric training program which showcased a 10cm improvement in DJ height following intervention of plyometric training on one side of the body and resistance training on the other side, showing a 1.3cm height improvement.[41] However, it should be mentioned this population differed to ours with utilising only males of a mean age of 22 ± 2 with no experience of regular resistance training. These results suggest this test provides value for assessing improvements in fast stretch-shortening cycle performance among ballet dancers. Moreover, as highlighted by Beattie and Flanagan, if the scores form athletes or dancers exceed that of the CV% calculated then the practitioner can be confident the change in DJ RSI is 'worthwhile' and is a result of a biological change in the athletes training status.[42]

For measures of jump height from the CMJ and SL CMJ, our findings suggest the between-session reliability was *excellent* (ICC = 0.95-0.98), with CV% ranging 3.0-5.0%. These findings are consistent with the literature[43,44,45], demonstrating the appropriateness of these tests for measuring strength performance utilising a slow stretch-shortening cycle in adolescent populations. This investigation is the first to determine these values in elite pre-professional ballet dancers. Notably, eight-week training interventions for both male[46] and female[47] adolescent athletes have demonstrated improvements in CMJ height that surpass the MDCs observed in this study. The measures of countermovement jump height appear to have sufficient reliability to detect changes after a relatively modest period of training (e.g., 1-2 training blocks). Although not statistically tested, our observation of the data aligned with Moir et al, suggesting no notable difference in reliability when using either the highest jump of three attempts or the mean of three attempts to calculate jump performance.[48] When deciding between using the highest jump or the mean of three attempts, practitioners should prioritise their philosophical approach rather than focusing exclusively on the accuracy of outcome measures. For instance, coaches evaluating a dancer's maximum force production capacity during a slow stretch-shortening cycle activity might select to analyse the highest jump as representative of CMJ performance.

For measures of maximal isometric force tests using the SL squat and SL PF test, these findings revealed *good* to *excellent* agreement (ICC = 0.87-0.98), with CV% ≤

4.8% for absolute vGRF and ≤ 4.9% for relative vGRF on both left and right limbs. This data is comparable to investigations measuring isometric strength qualities in an athletic population[49] and similar to that reported by Mattiussi et al, where ICC values ranged from 0.97 to 1.00, and CV% ranged from 2.0% to 5.9%.[14] However, it is important to acknowledge that, as the Brady et al. (2020) paper reviewed multiple studies, the participants varied in athletic ability, age, strength training experience, and joint angles compared to the dancer population in this study.[49] Furthermore, Mattiussi et al included both dancers and physically active males and females, with mean ages of 27.9 ± 6.3 and 29.3 ± 8.6, respectively.[14] This differs significantly from our study, which focused solely on dancers and involved a different age demographic. Notably, the MDC values in this study were higher than those reported by Kolokythas et al for the isometric mid-thigh pull (285–330N vs. 134N), suggesting that the isometric mid-thigh pull may offer greater sensitivity than the SL squat test.[15] Based on the MDC values presented in this investigation, maximal isometric force tests may not possess sufficient sensitivity to detect changes in strength following a relatively short strength training intervention. For example, Lynch et al found that recreational athletes following a 6-week bilateral or unilateral strength training programme, improved their bilateral and unilateral squat performance by 243N and 153N respectively.[50] These values fall below the MDC values observed in the present investigation's unilateral variant, representing 95% confidence intervals. Consequently, the isometric strength tests in the present study likely lack sufficient reliability to confidently detect performance changes after a single block (e.g., 4-6 weeks) of resistance training in adolescent ballet dancers. Therefore, detecting changes in maximal isometric force production during the SL squat may require extended training periods.

When examining the seated OHP test, this investigation revealed *excellent* between-session reliability for male dancers (ICC = 0.98, CV = 6.5%). To the authors' knowledge, no published research currently exists determining the reliability for the seated OHP to failure in healthy populations, with available research focusing predominantly on one repetition max testing in well trained men[51] or horizontal pressing movements.[52] However, assessment for strength endurance in the upper extremity demonstrate similar acceptable reliability. For example, Henriques-Neto and colleagues found the push-up test for maximum repetitions in young athletes between 9-18 years of age demonstrated *good* reliability (ICC = 0.86).[53] The OHP test was selected for this investigation due to its mechanical resemblance to lifts performed by male ballet dancers, involving significant shoulder elevation[54] that likely exceeds values observed during horizontal pressing activities.[55] Another consideration for the OHP test was that dancers were not restricted to performing lifts at a specific cadence, unlike in other tests of strength endurance.[56] This is an important consideration for practitioners using the OHP test, as research indicates that allowing individuals to choose their lifting tempo significantly increases the number of repetitions completed, average work performed, and average power displayed, compared to standardised cadences such as 2-second ascent with a 2-second descent, and a 2-second ascent with a 4-second descent.[57] In this study, lifting cadence was left uncontrolled to avoid the extended time needed for familiarisation and the difficulties in monitoring lifting speed, particularly when testing large cohorts with limited time available. Importantly, the data from this study show that the OHP test has sufficient sensitivity to detect potential in performance following an intervention.

Conclusion

The current study aimed to establish the between-session reliability of a testing battery assessing physical performance in elite adolescent ballet dancers. The data demonstrated *good* to *excellent* relative reliability for outcome measures related to jumping, lower extremity maximal isometric force production and upper extremity strength endurance tests. These results indicate that strength and power tests can be reliably integrated into a comprehensive performance testing battery to detect performance changes associated with strength gains following training interventions in this population. This expands testing options for adolescent ballet training centres and high-performance settings, ensuring confidence in their accuracy for measuring physical changes. The study suggests that these tests can effectively establish baseline performance data for power, strength and strength endurance, enabling practitioners to monitor performance changes accurately following physical interventions.

References

1. Shaw JW, Mattiussi AM, Brown DD, et al. Rehearsal and performance volume in professional ballet: a five-season cohort study. J Dance Med Sci. 2023;27(1):3-12. doi:10.1177/1089313x231174684
2. Twitchett EA, Koutedakis Y, Wyon MA. Physiological fitness and professional classical ballet performance: a brief review. J Strength Cond Res. 2009;23(9):2732-2740. doi:10.1519/JSC.0b013e3181bc1749
3. McCormack MC, Bird H, de Medici A, Haddad F, Simmonds J. The physical attributes most required in professional ballet: a delphi study. Sports Med Int Open. 2018;3(1):E1-E5. doi:10.1055/a-0798-3570
4. McGoldrick CD, Iacono AD, Morgan OJ, et al. Reliability, familiarization effect, and comparisons between a predetermined and a self-determined isometric-squat testing protocol. Int J Sports Physiol Perform. 2023;18(7):718-725. doi:10.1123/ijspp.2022-0480
5. Coogan SM, Hansen-Honeycutt J, Fauntroy V, Ambegaonkar JP. Upper-body strength endurance and power norms in healthy collegiate dancers: a 10-year prospective study. J Strength Cond Res. 2021;35(6):1599-1603. doi:10.1519/JSC.0000000000004016
6. Carroll KM, Wagle JP, Sole CJ, Stone MH. Intrasession and intersession reliability of countermovement jump testing in division-i volleyball athletes. J Strength Cond Res. 2019;33(11):2932-2935. doi:10.1519/JSC.0000000000003353
7. McMaster DT, Gill N, Cronin J, McGuigan M. A brief review of strength and ballistic assessment methodologies in sport. Sports Med. 2014;44(5). doi:10.1007/s40279-014-0145-2
8. Clemente FM, Badicu G, Hasan UC, et al. Validity and reliability of inertial measurement units for jump height estimations: a systematic review. Hum Mov. 2022;23(4):1-20. doi:10.5114/hm.2023.111548
9. Grgic J, Scapec B, Mikulic P, Pedisic Z. Test-retest reliability of isometric mid-thigh pull maximum strength assessment: a systematic review. Biol Sport. 2022;39(2):407-414. doi:10.5114/biolsport.2022.106149

10. Blagrove RC, Bishop C, Howatson G, Hayes PR. Inter-limb strength asymmetry in adolescent distance runners: Test-retest reliability and relationships with performance and running economy. J Sports Sci. 2021;39(3):312-321. doi:10.1080/02640414.2020.1820183
11. Harwood A, Campbell A, Hendry D, Ng L, Wild CY. Differences in lower limb biomechanics between ballet dancers and non-dancers during functional landing tasks. Phys Ther Sport. 2018;32:180-186. doi:10.1016/j.ptsp.2018.05.005
12. Orishimo KF, Liederbach M, Kremenic IJ, Hagins M, Pappas E. Comparison of landing biomechanics between male and female dancers and athletes, part 1. Am J Sports Med. 2014;42(5):1082-1088. doi:10.1177/0363546514523928
13. Nielsen J, Crone C, Hultborn H. H-reflexes are smaller in dancers from The Royal Danish Ballet than in well-trained athletes. Eur J Appl Physiol Occup Physiol. 1993;66(2):116-121. doi:10.1007/BF01427051
14. Mattiussi AM, Shaw JW, Cohen DD, et al. Reliability, variability, and minimal detectable change of bilateral and unilateral lower extremity isometric force tests. J Sport Exerc Sci. 2022;6(3):191-199. doi:10.36905/jses.2022.03.05
15. Kolokythas N, Metsios GS, Galloway SM, Allen N, Wyon MA. Reliability, variability and minimal detectable change of the isometric mid-thigh pull in adolescent dancers. J Dance Med Sci. 2024;28(1):14-20. doi:10.1177/1089313x231198421
16. Haff GG, Stone M, O'Bryant HS, et al. Force-time dependent characteristics of dynamic and isometric muscle actions. J Strength Cond Res. 1997;11(4):269-272. doi:10.1519/00124278-199711000-00014
17. Bishop C, Read P, Lake J, et al. Unilateral isometric squat: test reliability, interlimb asymmetries, and relationships with limb dominance. J Strength Cond Res. 2021;35:S144-S151. doi:10.1519/JSC.0000000000003079
18. Read P, Mc Auliffe S, Wilson MG, Myer GD. Better reporting standards are needed to enhance the quality of hop testing in the setting of ACL return to sport decisions: a narrative review. Br J Sports Med. 2021;55(1):23-29. doi:10.1136/bjsports-2019-101245
19. DeWolf A, McPherson A, Besong K, Hiller C, Docherty C. Quantitative measures utilized in determining pointe readiness in young ballet dancers. J Dance Med Sci. 2018;22(4):209-217. doi:10.12678/1089-313X.22.4.209

20. Maloney BM, Mattiussi AM, Cleather DJ, Price P, Shaw JW. Jumping demands during classical ballet class. Scand J Med Sci Sports. 2024;34(1). doi:10.1111/sms.14502
21. Twitchett E, Angioi M, Koutedakis Y, Wyon M. Video analysis of classical ballet performance. J Dance Med Sci. 2009;13(4):124-128. doi:10.1177/1089313x0901300405
22. Ramkumar PN, Farber J, Arnouk J, Varner KE, McCulloch PC. Injuries in a professional ballet dance company: a 10-year retrospective study. J Dance Med Sci. 2016;20(1):30-37. doi:10.12678/1089-313X.20.1.30
23. Teo W, Newton MJ, McGuigan MR. Circadian rhythms in exercise performance: implications for hormonal and muscular adaptation. J Sports Sci Med. 2011;10(4):600-606
24. Walter SD, Eliasziw M, Donner A. Sample size and optimal designs for reliability studies. Stat Med. 1998;17(1):101-110. doi:10.1002/(SICI)1097-0258(19980115)17:1<101::AID-SIM727>3.0.CO;2-E
25. Mirwald RL, Baxter-Jones ADG, Bailey DA, Beunen GP. An assessment of maturity from anthropometric measurements. Med Sci Sports Exerc. 2002;34(4):689-694. doi:10.1097/00005768-200204000-00020
26. Kozieł SM, Malina RM. Modified maturity offset prediction equations: validation in independent longitudinal samples of boys and girls. Sports Med. 2018;48(1):221-236. doi:10.1007/s40279-017-0750-y
27. Moran J, Hammami R, Butson J, et al. Do verbal coaching cues and analogies affect motor skill performance in youth populations? PLoS One. 2023;18(3). doi:10.1371/journal.pone.0280201
28. Oliver JL, Barillas SR, Lloyd RS, Moore I, Pedley J. External cueing influences drop jump performance in trained young soccer players. J Strength Cond Res. 2021;35(6):1700-1706. doi:10.1519/JSC.0000000000002935
29. Morin M, Duchesne E, Bernier J, Blanchette P, Langlois D, Hébert LJ. What is known about muscle strength reference values for adults measured by hand-held dynamometry: a scoping review. Arch Rehabil Res Clin Transl. 2021;4(1):100172. doi:10.1016/j.arrct.2021.100172
30. Landin D, Thompson M, Reid M. Knee and ankle joint angles influence the plantarflexion torque of the gastrocnemius. J Clin Med Res. 2015;7(8):602-606. doi:10.14740/jocmr2107w

31. Earp JE, Kraemer WJ, Newton RU, et al. Lower-body muscle structure and its role in jump performance during squat, countermovement, and depth drop jumps. J Strength Cond Res. 2010;24(3):722-729. doi:10.1519/jsc.0b013e3181d32c04
32. Haff GG. Strength – isometric and dynamic testing. In: Performance Assessment in Strength and Conditioning. Vol 1. 1st ed. Routledge; 2018:166-192
33. Atkinson G, Nevill AM. Statistical methods for assessing measurement error (reliability) in variables relevant to sports medicine. Sports Med. 1998;26(4):217-238. doi:10.2165/00007256-199826040-00002
34. Koo TK, Li MY. A guideline of selecting and reporting intraclass correlation coefficients for reliability research. J Chiropr Med. 2016;15(2):155-163. doi:10.1016/j.jcm.2016.02.012
35. Riemann BL, Lininger MR. Principles of statistics: what the sports medicine professional needs to know. Clin Sports Med. 2018;37(3):375-386. doi:10.1016/j.csm.2018.03.004
36. Xu J, Turner A, Comfort P, et al. A systematic review of the different calculation methods for measuring jump height during the countermovement and drop jump tests. Sports Med. 2023;53(5):1055-1072. doi:10.1007/s40279-023-01828-x
37. Azevedo AM, Oliveira R, Vaz JR, Cortes N. Oxford foot model kinematics in landings: a comparison between professional dancers and non-dancers. J Sci Med Sport. 2020;23(4):347-352. doi:10.1016/j.jsams.2019.10.017
38. Guy-Cherry D, Alanazi A, Miller L, Staloch D, Ortiz-Rodriguez A. Landing styles influences reactive strength index without increasing risk for injury. Sports Med Int Open. 2018;2(2):E35-E40. doi:10.1055/a-0608-4280
39. Blanco P, Nimphius S, Seitz LB, Spiteri T, Haff GG. Countermovement jump and drop jump performances are related to grand jeté leap performance in dancers with different skill levels. J Strength Cond Res. 2021;35(12):3386-3393. doi:10.1519/JSC.0000000000003315
40. Comyns TM, Murphy J, O'Leary D. Reliability, usefulness, and validity of field-based vertical jump measuring devices. J Strength Cond Res. 2023;37(8):1594-1599. doi:10.1519/JSC.0000000000004436

41. Kubo K, Morimoto M, Komuro T, et al. Effects of plyometric and weight training on muscle-tendon complex and jump performance. Med Sci Sports Exerc. 2007;39(10):1801-1810. doi:10.1249/mss.0b013e31813e630a
42. Beattie K, Flanagan EP. Establishing the reliability & meaningful change of the drop-jump reactive-strength index. J Aust Strength Cond. 2015;23(5):12-18
43. Thomas C, Dos'Santos T, Comfort P, Jones PA. Between-session reliability of common strength- and power-related measures in adolescent athletes. Sports (Basel). 2017;5(1):15. doi:10.3390/sports5010015
44. Haines BR, Bourdon PC, Deakin G. Reliability of common neuromuscular performance tests in adolescent athletes. J Aust Strength Cond. 2016;24(4):16-22. doi:11541.2/135298
45. Snyder BW, Munford SN, Connaboy C, Lamont HS, Davis SE, Moir GL. Assessing plyometric ability during vertical jumps performed by adults and adolescents. Sports (Basel). 2018;6(4):132. doi:10.3390/sports6040132
46. Chelly MS, Hermassi S, Aouadi R, Shephard RJ. Effects of 8-week in-season plyometric training on upper and lower limb performance of elite adolescent handball players. J Strength Cond Res. 2014;28(5):1401-1410. doi:10.1519/jsc.0000000000000279
47. Ozbar N, Ates S, Agopyan A. The effect of 8-week plyometric training on leg power, jump and sprint performance in female soccer players. J Strength Cond Res. 2014;28(10):2888-2894. doi:10.1519/JSC.0000000000000541
48. Moir G, Shastri P, Connaboy C. Intersession reliability of vertical jump height in women and men. J Strength Cond Res. 2008;22(6):1779-1784. doi:10.1519/JSC.0b013e318185f0df
49. Brady CJ, Harrison AJ, Comyns TM. A review of the reliability of biomechanical variables produced during the isometric mid-thigh pull and isometric squat and the reporting of normative data. Sports Biomech. 2020;19(1):1-25. doi:10.1080/14763141.2018.1452968
50. Lynch AE, Davies RW, Allardyce JM, Carson BP. The effect of unilateral versus bilateral strength training on isometric-squat peak force and interlimb asymmetry in young, recreationally strength-trained men. Int J Sports Physiol Perform. 2023;18(2):195-203. doi:10.1123/ijspp.2022-0299

51. Soriano MA, García-Ramos A, Torres-González A, et al. Validity and reliability of a standardized protocol for assessing the one repetition maximum performance during overhead pressing exercises. J Strength Cond Res. 2021;35(11):2988-2992. doi:10.1519/JSC.0000000000003284
52. Eckel TL, Watkins CM, Archer DC, et al. Bench press and pushup repetitions to failure with equated load. Int J Sports Sci Coach. 2017;12(5):647-652. doi:10.1177/1747954117733879
53. Henriques-Neto D, Minderico C, Peralta M, Marques A, Sardinha LB. Test-retest reliability of physical fitness tests among young athletes: The FITescola® battery. Clin Physiol Funct Imaging. 2020;40(3):173-182. doi:10.1111/cpf.12624
54. McKean MR, Burkett BJ. Overhead shoulder press - in-front of the head or behind the head? J Sport Health Sci. 2015;4(3):250-257. doi:10.1016/j.jshs.2013.11.007
55. Suprak DN, Bohannon J, Morales G, Stroschein J, San Juan JG. Scapular kinematics and shoulder elevation in a traditional push-up. J Athl Train. 2013;48(6):826-835. doi:10.4085/1062-6050-48.5.08
56. Artero EG, España-Romero V, Castro-Piñero J, et al. Reliability of field-based fitness tests in youth. Int J Sports Med. 2011;32(3):159-169. doi:10.1055/s-0030-1268488
57. LaChance PF, Hortobagyi T. Influence of cadence on muscular performance during push-up and pull-up exercise. J Strength Cond Res. 1994;8(2):76-79. doi:10.1519/00124278-199405000-00003

List of Legends

Figure 1. Timeline for data collection across the 12-week testing period.