Reliability of a standardized protocol of the Single Leg Heel Rise Test

Anna Schrefl*[1], Nico Kolokythas[2], Matteo Stamm[1], Daniel Erlacher[1], Andrea Schärli[1]

[1] Institute of Sport Science, University of Bern, Bern, Switzerland
[2] Research and Performance Department, Elmhurst Ballet School, Birmingham, United Kingdom

* anna.schrefl@students.unibe.ch

ORIGINAL ARTICLE

Submitted: December 30, 2023

Accepted: June 24, 2024

Published: October 24, 2024

Editor-in-Chief:

Claudio R. Nigg, University of Bern, Switzerland

Section Editor:

Martin Keller, University of Basel, Switzerland

ABSTRACT

The single leg heel rise (SLHR) test is a widely used method for assessing calf muscle-tendon unit endurance in various fields, including medicine, sports, and dance. The objectives of this study were to examine the reliability of a standardized SLHR test protocol with a heel rise measurement device and to investigate the relationship between SLHR repetitions, SLHR total work, and both maximal voluntary isometric contraction (MVIC) and reactive strength outcomes of the plantar flexors. Twenty-one sports students (8 females, 13 males) were assessed for SLHR outcomes (number of repetitions, height of heel rises, total positive work performed) in two data collection sessions, as well as unilateral lower extremity MVIC and reactive strength. Test-retest results showed excellent reliability and low variability for the number of repetitions (Intraclass correlation coefficient (ICC) = 0.91, 95% CI: 0.75 to 0.97; coefficient of variability (CV): 8.1–8.7%), height of the heel rises (ICC = 0.93, 95% CI: 0.77 to 0.98; CV: 3.8–5.7%) and total positive work performed (ICC = 0.96, 95% CI: 0.86 to 0.99; CV: 6.2–10.2%) in both feet. No significant correlation was identified between SLHR repetitions, MVIC, and reactive strength outcome measures. A moderate correlation was observed between the total positive work performed in the SLHR and reactive strength outcomes that could be explained by the participants' body weight. The SLHR test provides reliable measures for lower leg muscular endurance, yet it does not predict plantar flexor maximal strength or reactive strength. In future studies, we advise employing this standardized protocol for screenings of athletes and dancers.

Keywords

single leg heel rise, calf raise, muscular endurance, plantar flexors, calf performance

Citation:

Schrefl, A., Kolokythas, N., Stamm, M., Erlacher, D., & Schärli, A. (2024). Reliability of a standardized protocol of the Single Leg Heel Rise Test. *Current Issues in Sport Science*, *9*, Article 009. https://doi.org/10.36950/2024.9ciss009

Introduction

Over the past decades, physical screening tests have become increasingly popular for assessing fitness levels, identifying muscle imbalances, deficits in movements, and muscle weakness (Bonazza et al., 2017; Chorba et al., 2010; Green et al., 2022; Yeung et al., 2009). As such, screening protocols have been widely endorsed by clinicians, researchers, and trainers as an important component of injury prevention strategies across various populations, including general (De la Motte et al., 2017), athletic (McKeown et al., 2014), and dance (Armstrong & Relph, 2021; Gamboa et al., 2008) communities. Apart from injury prevention, field-based screening protocols are also used to quantify training progress and skill acquisition in recreational and professional athletes (Rogers et al., 2019) and dancers (Liederbach, 1997). Next to screening programs that evaluate movement competency, various protocols to assess an individual's flexibility, agility, power, strength, and muscular endurance are used in the field (Bird & Markwick, 2016; Dennis et al., 2008; Evans et al., 2007). When it comes to measuring the muscular endurance of the calf muscle-tendon unit (MTU), the single leg heel rise (SLHR) test has been established as a commonly used test in medicine (Fernandez et al., 2023; Hébert-Losier et al., 2009), sports (Hébert-Losier et al., 2009, 2023), and dance (Schrefl et al., 2023; Thomas, 2003). This test is recommended to monitor the plantar flexor performance of athletes (Hébert-Losier et al., 2023) and dancers (Thomas, 2003), to incorporate it into athletes' rehabilitation programs (Green et al., 2022), and to assess the pointe readiness of young ballet dancers (DeWolf et al., 2018).

Generally, the SLHR test is assessed by counting the total number of heel rises completed while standing on one foot, with a person's body weight serving as resistance (Hébert-Losier et al., 2017). Although practical, there is a lack of normative data available for athletes and dancers. This knowledge gap arises from the different variations of test protocols used in sports and dance research, making it challenging to compare outcomes (Hébert-Losier et al., 2009; Schrefl et al., 2023). However, standardizing the SLHR test protocol is crucial among professional athletes, dancers, and dance students to establish norms necessary for improving training and enhancing rehabilitation protocols after injuries.

Two critical factors that can significantly influence test results are the execution form and the height of the heel rises (Hébert-Losier et al., 2009; Thomas, 2003). This is especially important in aesthetic sports and dance, where precise technique, alignment, and achieving a full range of plantar flexion are essential for training and performance. Thus, measuring and controlling the height of the heel rises is key. This approach allows for a thorough assessment of the number of repetitions and quantifies the total work performed (Hébert-Losier et al., 2009). Hébert-Losier et al. (2023) demonstrated a moderate, inverse correlation (r = -.55, p < .05) between the total work completed during the SLHR test and the 10-meter sprint times among Rugby Union players. However, they highlighted that the number of repetitions did not show any statistically significant associations (r = -.28, p = .29) with the 10-meter sprint times. Their argument emphasized that solely considering the number of repetitions overlooks the overall range or distance covered during the SLHR test, thus inadequately reflecting the total work completed (Hébert-Losier et al., 2023).

To measure the precise height of heel rises, 3-D motion capture systems are considered the gold standard (Pires et al., 2020). For practical field tests, visual encoders and, more recently, a mobile app designed for calf raises are utilized due to their feasibility (Fernandez et al., 2023). However, these devices do not control the height of the heel rises during the SLHR test. In dance, devices that use simple rubber-band systems to control the height of each heel rise are implemented in screening and testing settings (DeWolf et al., 2018; Schrefl et al., 2023; Sman et al., 2014). As these devices, in turn, do not measure the exact height of the heel rises, they are not suitable for calculating the total work performed. Thus, a device for the SLHR test should be able to measure and control the height of the heel rises while also being easily accessible for field testing.

Hébert-Losier et al. (2023) identified correlations between the SLHR test and the ability to accelerate over a short distance, while Sara et al. (2021) argued that this test inadequately measures plantar flexor maximal strength. However, Sara et al. (2021) solely focused on the number of repetitions and did not consider the total work completed. The SLHR test, a dynamic task performed under constant submaximal load until failure, raises questions about its efficacy in assessing maximal and reactive strength. It is crucial to understand the capabilities and limitations of a test when incorporating it into a screening setting. Given the widespread use of the SLHR test, its suitability for evaluating plantar flexor strength warrants clarification. Further investigation into the relationship between the number of repetitions, total work completed, and these parameters is imperative, as both maximal and reactive strength are essential parameters for jumping performance in sports and dance.

Therefore, the aim of this study was I) to evaluate the test-retest and interrater reliability of a standardized SLHR test protocol using a simple and new device to control and measure the height of the heel rises, and II) to assess the associations between SLHR repetitions, SLHR total work, and maximal voluntary isometric contraction (MVIC) as well as reactive strength outcomes.

Methods and Materials

Study design

A repeated-measures reliability study was conducted. Participants attended two data collection sessions with an interval of 12 to 14 days between the sessions. An a priori power analysis was conducted using the Bonett formula (Borg et al., 2022) for calculating sample sizes for reliability studies. The significance level was set at .05, and the true ICC was anticipated to be high (.95) (Ross & Fontenot, 2000). The minimum sample size was calculated to be 21 participants. The dropout rate was expected to be 10%. Thus, a sample size of n = 24 was adequate for this study.

Participants

Twenty-four healthy sports students (9 females, 15 males) volunteered to participate in the study. Exclusion criteria were any current injury of the lower extremities. Written informed consent was obtained after all participants were given details about the procedures and the associated risks of the tests. The study protocol was designed in accordance with the Declaration of Helsinki by the ethics committee of the local Faculty of Human Sciences (Approval Nr. 2022-11-04).

Measurements and procedure

The first data collection was used to calculate the reliability of the SLHR test. The measurements included the SLHR test and MVIC of the plantar flexors with a maximal isometric strength test. Test-retest and interrater reliability of the SLHR test and the minimal detectable change (MDC) were calculated. Prior to the first data collection, participants completed a health, demographic, and leg-dominance questionnaire. To determine leg dominance, participants were asked which leg they would use to shoot a ball at a target (van Melick et al., 2017). The testing session started with a standardized 5-minute warm-up, including three minutes on a treadmill of fast walking at a self-selected pace and two minutes of global physical mobilization exercises, particularly the ankle joint. The maximal isometric strength test was performed on each side, followed by a 15-minute break. During the break, a marker was set on the lateral malleoli with a skin-friendly pen. Participants then performed the SLHR test on each side. In the second data collection, reactive strength scores of the lower extremities were calculated with a 30 cm drop jump (DJ) test. The second data collection session started with the same standardized warm-up. Participants then performed the DJ test, followed by a 7-minute break and the SLHR test on each side.

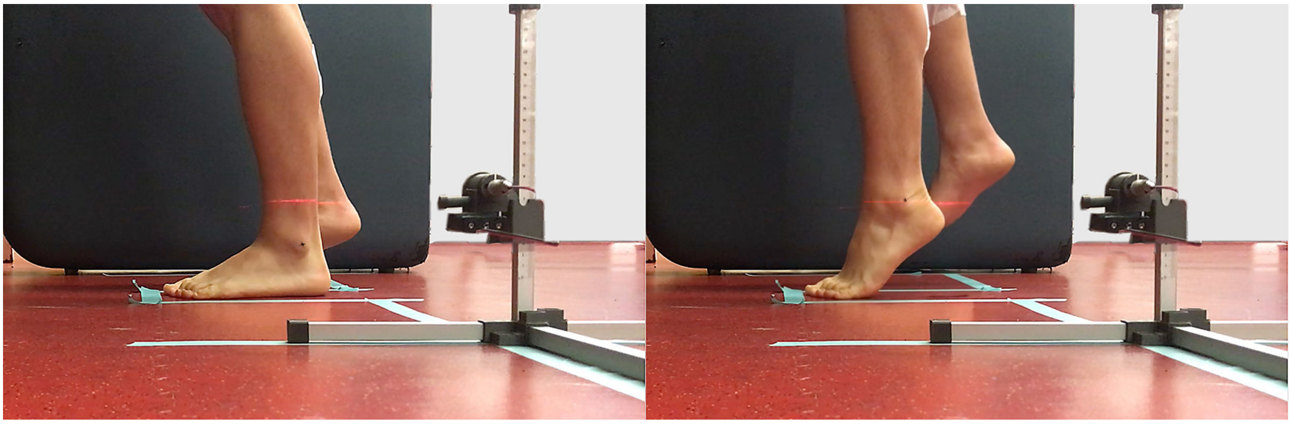
SLHR test

The protocol was based on previous studies using this test (Hébert-Losier et al., 2009; Lunsford & Perry, 1995; Thomas, 2003). The participant stood in front of a wall bar in a single-leg stance and placed two fingers of each hand on a bar to use for balance. A digital metronome (Seiko SQ-60) was set at 60 beats per minute (BPM). The examiner gave a three-count preparation. The participant lifted the heel on one beat and returned the heel to the floor on the next beat. The participant was instructed to always raise to the highest heel rise with a straight leg, to keep the non-working leg off the ground at all times, to keep the foot centered, and to avoid leaning against the bar.

The participant first performed a try-out of two SLHRs on the dominant leg on the BPM of the metronome. Afterward, the light beam of a custom-made laser pointer (LP) device was aligned with the marker of the lateral malleolus of the dominant leg to measure the height of the heel rises. Initially, the light beam was set at the marker while the participant was standing, establishing the baseline height. The participant then performed a maximal heel rise, and the light beam was adjusted to this position. The height was measured in millimeters on a scale on the LP device, and the light beam was used to indicate the maximal heel rise height during the test (Figure 1). After a short break of approximately two minutes, the test started on the dominant leg. The participant performed as many heel rises as possible. The test was stopped either by the participant or by the examiner. Termination criteria were the following: not reaching maximal heel rise height within a range of 5 mm, bending the knee of the standing leg, or losing the alignment, such as leaning towards the bar, for two consecutive repetitions. After a short break of three minutes, the test was repeated on the other leg. The last repetition was counted as valid if 75 % of the maximal heel rise height was reached.

Maximal isometric strength test

To measure the MVIC of the plantar flexors, the participant stood on a mobile 3D force plate (Kistler Type 9260AA, Kistler Group, Switzerland) positioned under a barbell rack. The rack was fixated with a total weight of 120 kg. For each participant, the distance between the rack and the force plate was adjusted according to the participant's height by adding or removing 5 cm diameter wooden plates underneath the force plate. We then secured the force plate to a stable surface and calibrated it before each participant to ensure the accuracy of each test. The barbell was positioned on the upper trapezius of the participants in a high-bar back squat stance (Mattiussi et al., 2022). Then the participant shifted the weight onto one leg and lifted the heel of the foot to a comfortable height. The other foot was not touching the ground. Verbal instructions were given to the participant to keep the knee slightly bent and to keep the pelvis in a neutral position. The participant then applied maximum pressure against the barbell for three seconds. The participant was instructed to press with full strength from the leg by pushing the ball of the standing foot against the force plate without bending or extending the knee or lifting the shoulders. The test was performed on each leg three times, starting with the dominant leg, with a one-minute break between the trials. A fourth repetition was added when a clear plateau in the force reading was not reached, or no data had been recorded for technical reasons.



A. Starting position.

B. Heel rise. The beam of the laser pointer was set at the height of the heel rise.

Figure 1 *Custom-made laser pointer (LP) device to control the height of the heel rises. The beam of the LP was set on the marker of the lateral malleolus. The height was displayed in mm on a scale on the LP device.*

Drop jump test

Reactive strength was measured with the DJ test (Möck et al., 2023). In order to be physically prepared for this test, each participant executed ten consecutive small explosive jumps with minimal contact time. These jumps were added at the end of the standardized warm-up of the second test session. The participant then stood upright on a 30 cm high platform with the hands fixed on the waist. To initiate the jump, the participant "stepped out" from the platform with one foot and dropped onto a force plate (Kistler Type 9290DD, Kistler Group, Switzerland), aiming to jump as high and fast as possible. All participants were verbally instructed to jump explosively, to keep the contact time on the ground as short as possible, and to pay attention to a bouncing jump execution (Möck et al., 2023). The DJ test was performed twice with a pause of 45 seconds in between the jumps. A third DJ was performed when the contact time of the first two jumps exceeded 250 milliseconds, or the participant stepped forward after landing (Möck et al., 2023).

Outcome measures

During the SLHR test, the number of repetitions (Reps) was counted by one examiner and the height of the heel rises was recorded by a second examiner. Additionally, videos of both legs were recorded with a fixed mobile phone camera, in which only the participant's leg could be seen. Two independent blinded raters counted the number of repetitions on the video recordings. The total positive work of each trial was calculated by multiplying the body weight (BW) by the total distance traveled (Zellers et al., 2017).

$$Work_{positive}(Joule) = BW(N) * distance(m)$$

Furthermore, plantar flexor total MVIC and MVIC relative to BW were measured during the maximal isometric strength test. Ground contact time (GCT) and jump height (JH) values were collected in the DJ test. The reactive strength index (RSI) of the jumps was then calculated by dividing the JH by GCT and is expressed as m/s units (Dello Iacono et al., 2016).

Statistical analysis

The means and standard deviation (*SD*) were calculated for each variable (Table 2). Test-retest and inter-rater intraclass correlation coefficient (*ICC*) estimates and their 95% confidence interval (*CI*) were based on a 2-way mixed-effects model with an absolute agreement (Koo & Li, 2016). Test-retest reliability was calculated for the number of repetitions, the height of the heel rises, and total positive work performed. To differentiate training-related changes from noise, the standard error of the measurement (*SEM*), the minimal detectable change (*MDC*), and the coefficient of variation (*CV*) were calculated to establish random error scores between the first and second data collection (Comfort et al., 2015; Howarth et al., 2022). *SEM* was calculated using the *SD* of the differences in scores (SD_{diff}) (Hopkins, 2000):

$$SEM = SDdiff \times \sqrt{1 - ICC}$$

MDC was calculated using the formula

$$MDC = 1.96 \times \sqrt{2} \times SEM.$$

The following formula was used to calculate the *CV* (Möller et al., 2005):

$$CV = 100 \times s_w / mean_{total}$$

and

$$s_w = \sqrt{\sum (SD^2_{diff}/2n)}.$$

Paired samples t-tests were conducted to detect the differences in the sessions' mean. The statistical significance was set at $p < .05$. Moreover, Bland-Altman plots were created to evaluate if there was bias in the measures and linear regressions were performed to detect proportional bias.

A correlation matrix was then performed to determine the relation between the total work performed and the number of Reps of the SLHR test and plantar flexor total MVIC, MVIC relative to BW, GCT, as well as RSI of the DJ tests. The calculations were executed using IBM SPSS Statistics (version 29). Reactive strength outcome measurements were processed using Kistler MARS software (version 2019), and all MVIC data were processed using MATLAB (Version R2022a; Math Works Inc., Natick, MA, USA).

Results

Twenty-four participants were enrolled in the study. All participants completed the first test session. Only twenty-two participants attended the second test session due to illness and scheduling problems. The data of one participant had to be removed as the participant had a strenuous competition the night before the second test session. In total, the data from twenty-one participants were used for further calculations (12.5% dropout rate). The characteristics of the participants are shown in Table 1. Male and female participants showed similar results in all parameters of the SLHR test, plantar flexor MVIC relative to BW, and GCT. Males showed higher values of total plantar flexor MVIC, RSI, and jump height in the DJs. A summary of the results for all test variables is displayed in Tables 2 and 3.

Interrater ICC of the SLHR test

Three independent raters counted the number of repetitions for both test sessions of the SLHR test, with one rater counting the repetitions live and two raters watching the recorded videos. Excellent interrater reliability was found for both test sessions (ICC 1.0, 95% CI: 0.99 to 1.0).

Table 1
Basic characteristics of 21 participants.

Sex (females/males)	Age (years)	Weight (kg)	Height (cm)	Poly athletic training experience (years)
8 / 13	22.6 ± 1.7	68.8 ± 10.1	174.4 ± 8.9	10.5 ± 4.9

Note. Values are provided as mean ± standard deviation.

Table 2
Descriptive data of the Maximal Strength and Drop Jump Test.

		Drop Jump Test		
	Sex	**GCT (s)**	**Jump height (m)**	**RSI (m/s)**
Both Feet	Female	0.17 ± 0.016	0.17 ± 0.03	1.02 ± 0.17
	Male	0.19 ± 0.016	0.25 ± 0.04	1.38 ± 0.26

		Maximal Strength Test	
	Sex	**MVIC (N)**	**rMVIC (N×kg^{-1})**
Dominant side	Female	735 ± 139	12.7 ± 1.2
	Male	902 ± 138	12.2 ± 2.2
Non-dominant side	Female	726 ± 112	12.0 ± 1.3
	Male	891 ± 169	12.0 ± 2.7

Note. Values are provided as mean ± standard deviation. MVIC = maximal voluntary isometric contraction, rMVIC = MVIC relative to BW, BW = body weight, DJ = drop jump, GCT = ground contact time, RSI = reactive strength index.

Table 3
Reliability of the SLHR Test.

		Test		Retest		Statistics				
	Side	***Mean***	***SD***	***Mean***	***SD***	***SD$_{diff}$***	***SEM***	***MDC (%)***	***CV***	***ICC (95% CI)***
Number of Repetition (n)	Dominant	36.2	7.2	36.2	7.6	4.6	1.45	4.01 (11.1)	8.73	0.90 (0.75-0.96)
	Non-dominant	35.8	7.1	36.8	8.5	4.1	1.14	3.16 (8.7)	8.10	0.92 (0.82-0.97)
Height of heel rises (mm)	Dominant	67.3	8.9	68.6	9.7	5.4	1.67	4.62 (6.8)	5.67	0.91 (0.77-0.96)
	Non-dominant	67.9	8.3	66.3	7.5	3.4	0.80	2.22 (3.3)	3.81	0.94 (0.85-0.98)
Total work (Joule)	Dominant	1651.4	514.8	1680.8	528.4	244.8	57.92	160.56 (9.7)	10.21	0.94 (0.86-0.98)
	Non-dominant	1636.8	444.9	1648.1	486.2	148.2	23.44	64.96 (4.0)	6.25	0.98 (0.94-0.99)

Note. *SD* = standard deviation, *SDdiff* = SD of differences, *SEM* = standard error of measurement, *MDC* = minimal detectible change, *CV* = coefficient of variation, *ICC* = Intraclass Correlation Coefficient. *ICC* values < 0.50 indicate poor reliability, values between 0.50 and 0.75 demonstrate moderate reliability, values between 0.75 and 0.90 indicate good reliability, and values > 0.90 suggest excellent reliability.

Test-retest ICC of the SLHR test

The test-retest ICC values for the number of repetitions of the SLHR test indicate good (0.90) to excellent (0.92) reliability for the dominant and non-dominant foot, respectively. Test-retest reliability of the height of the heel rises and total positive work performed were excellent (0.91–0.94) for both feet. Paired samples t-tests found no significant differences between the test and retest trials. Values for the *Mean*, *SD*, *SEM*, *MDC*, *CV*, and *ICC* values of the two test sessions are displayed in Table 3.

Means and standard deviations of the differences (SD_{diff}) between the test and retest of the SLHR were normally distributed. The Bland-Altman plots for test-retest reliability for the number of repetitions, height of heel rises, and total positive work performed are shown in Figure 2. No proportional bias was detected.

Correlations

No significant correlations were found between the number of SLHR repetitions and various factors such as total MVIC, relative MVIC to BW (rMVIC), jump height, GCT, and RSI. Additionally, the total work of the SLHR test did not significantly correlate with total MVIC, rMVIC, or GCT. Moderate correlations were observed between the SLHR test's total positive work performed and both jump height and RSI in both feet (Table 4).

Table 4

Correlations between SLHR repetitions, total work, and maximal and reactive strength parameters

	MVIC	rMVIC	JH	GCT	RSI
Reps – Dominant side	-.12	-.31	.20	.08	.16
Reps – Non-dominant side	-.05	-.15	.21	.01	.20
Total work – Dominant side	.26	-.30	.49*	.11	.46*
Total work – Non-dominant side	.40	-.13	.49*	.11	.47*

MVIC = maximal voluntary isometric contraction, rMVIC = MVIC relative to BW, JH = jump height, GCT = ground contact time, RSI = reactive strength index.

*$p < .05$.

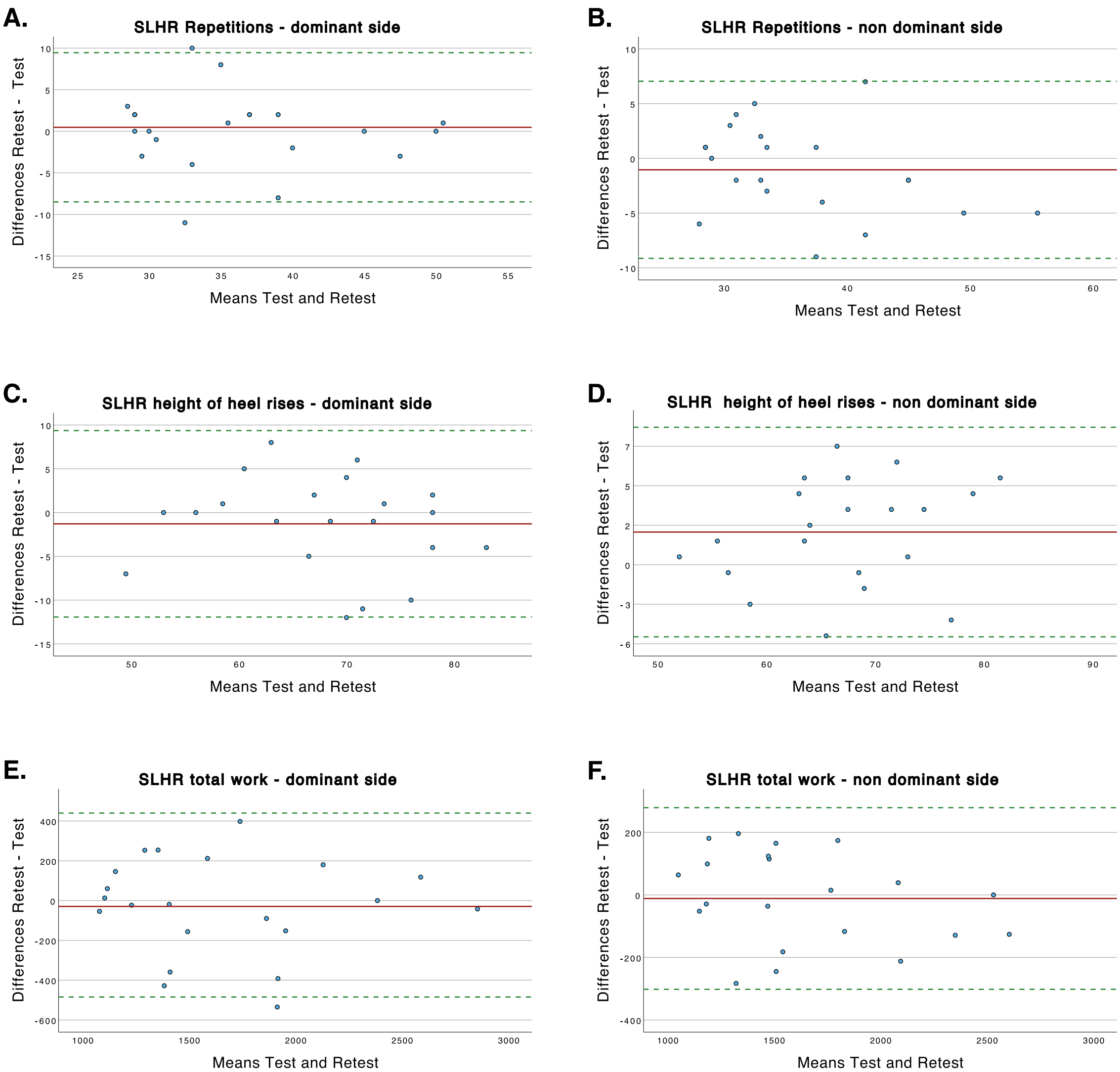


Figure 2 *Bland-Altman plots of the SLHR test. (A) Repetitions of the dominant foot. (B) Repetitions of the non-dominant foot. (C) Height of heel rises of the dominant foot. (D) Height of heel rises of the non-dominant foot. (E) Total positive work performed of the dominant foot. (F) Total positive work performed of the non-dominant foot. The red lines show the means of the differences (= bias) between the two methods, and the dotted horizontal green lines show the upper and lower 95% limits of agreement (= bias ± 1.96 × SD).*

Discussion

This study aimed to assess the reliability of a standardized test protocol for the SLHR test and explore the associations between SLHR repetitions, SLHR total work performed, and outcomes related to both maximal strength and reactive strength. Our results demonstrated excellent interrater ICC of the SLHR test protocol as well as good to excellent test-retest ICC values for the number of repetitions, the height of the heel rises, and total work performed. Furthermore, our findings showed no significant correlations between the number of repetitions in the SLHR test and both the outcomes of the plantar flexor MVIC and reactive strength test measured by the drop jump test. SLHR total work performed was moderately correlated with RSI and jump height in the reactive strength measurements.

The test-retest ICC estimates of SLHR repetitions align with previous studies (Haber et al., 2004; Nunes et al., 2019; Rowley et al., 2015), with the 95% CI ranging from moderate (0.75) to excellent (0.97) and good (0.77) to excellent (0.98) for SLHR repetitions and SLHR height of the heel rises, respectively. Bland-Altman plots did not reveal any proportional bias, suggesting that factors like a general learning effect did not impact the retest measurements. For this population, the meaningful change in test performance, known as the MDC, was in the range of 3 to 4 repetitions or 9 to 11 percent of an individual's score. The MDC represents a significant shift in test results that cannot be attributed to chance, making it highly relevant for practitioners. The variability of the within-subject differences, as indicated by the CV, was considered good for both the number of repetitions (<9%) and the height of the heel rises (<6%) (Hopkins, 2000; Shechtman, 2013).

While the reliability of the SLHR test was good to excellent, and no proportional bias was found, a few noteworthy observations from the test and retest trial sessions could be of interest. Participants were instructed to place two fingers lightly on a bar for balance support. However, several participants needed correction to avoid pushing from the fingers during the end of the trials. Therefore, we suggest positioning the fingers beneath the bar to minimize the likelihood of using the hands for additional support.

Another observation was the location of muscle fatigue. Following the test, participants were questioned about the primary site of their muscular fatigue. While most indicated the calf muscles, a few highlighted the gluteal muscles instead. This pertains to a general concern with the SLHR test. According to Sara et al. (2021), this test does not solely target the plantar flexors but assesses the fatiguability of the whole lower limb. Other muscles in the lower extremities also contribute to its execution (Sara et al., 2021). Additionally, the core musculature needs to stabilize the trunk to maintain an upright position while performing the task. Overall, the SLHR test is most accurately characterized as evaluating lower leg muscular endurance via a functional, dynamic task (Sara et al., 2021). The assessment does not evaluate calf MTU in isolation, meaning that weaknesses in other muscle groups can affect the performance of the test. However, in many dance forms, where heel rises are part of training and performances, a functional test of lower leg muscular endurance is warranted.

No correlation was observed between either the number of repetitions or the total positive work performed in the SLHR test and the MVIC relative to body weight. This lack of correlation was expected, as maximal strength and the capacity to sustain a task over a duration require distinct physiological mechanisms (Hunter et al., 2004; Sara et al., 2021). During dynamic fatiguing contractions, the muscle's capacity to adapt metabolically to fatigue and maintain neural drive are the primary physiological factors, rather than its cross-sectional area (Sara et al., 2021). We observed a moderate correlation between SLHR total work and the jump height as well as the RSI in the reactive strength test. However, these correlations might have been influenced by the participants' body weight acting as a confounding factor. When excluding body weight from the calculations, no significant correlations were found. Therefore, our results suggest that the SLHR test is a poor predictor for maximal strength and jump performance in screenings, highlighting its limitation in assessing athletes' and dancers' jumping performance. Incorporating an additional test for maximal plantar flexor strength is recommended for a more comprehensive assessment of the calf MTU performance.

In terms of the number of repetitions, the average achieved was 36, spanning from 27 to 58. This is in line with the study of Héber-Losier et al. (2017). The differences in sport types and activity levels of the participants may explain the wide range from 27 to 58 repetitions in our population. No differences were found between males and females in this study. Male and female participants performed more repetitions on their dominant side and showed higher heel rise height on their non-dominant side. However, the differences were minor and can be considered normal according to the asymmetry formula of MacSweeny et al. (2024).

Some limitations of our study should be pointed out. The time span between the test and retest ranged from 12 to 14 days, which was relatively long and could potentially impact the results. However, the outcome did not reveal any proportional bias. Also, although participants were instructed to refrain from strenuous activities 24 hours before the test trials, it is important to note that they were sports students. As such, preceding activities could have influenced the performance of the SLHR test. Furthermore, a force plate for each leg in the drop jump test would have given a more comprehensive insight into each leg, leading to more informative results.

Moreover, our calculations of the total work during SLHR repetitions may have a certain lack of accuracy due to the allowance of a 5 mm range for participants to achieve their maximal heel rise height, which was predetermined before the test. The LP device has demonstrated great practicability in screening settings involving numerous athletes or dancers. Its ability to swiftly determine the maximal heel rise height, along with the capability to control heel rise height during the test, are notable advantages. Nevertheless, further investigation is warranted to validate the LP device against gold standard methods like 3-D motion capture systems or established measurements such as the calf raise app in future studies.

Conclusion

Our study affirms the reliability of the SLHR test when executed within a strict and standardized test protocol. Within this standardized test protocol, we observed good to excellent test-retest and interrater intraclass correlation coefficients (ICC). We, therefore, suggest implementing this protocol in sports and dance screenings, as well as in future studies, to establish normative and comparable values for professional athletes, dancers, and dance students. While the SLHR test assesses lower leg muscular endurance, it is a poor predictor of jump performance. This factor should be considered when designing screening programs for athletes and dancers.

References

Armstrong, R., & Relph, N. (2021). Screening Tools as a Predictor of Injury in Gymnastics: Systematic Literature Review. *Sports Medicine - Open*, *7*(1), 73. https://doi.org/10.1186/s40798-021-00361-3

Bird, S. P., & Markwick, W. J. (2016). Musculoskeletal Screening and Functional Testing: Considerations for Basketball Athletes. *International Journal of Sports Physical Therapy*, *11*(5), 784–802.

Bonazza, N. A., Smuin, D., Onks, C. A., Silvis, M. L., & Dhawan, A. (2017). Reliability, Validity, and Injury Predictive Value of the Functional Movement Screen: A Systematic Review and Meta-analysis. *The American Journal of Sports Medicine*, *45*(3), 725–732. https://doi.org/10.1177/0363546516641937

Borg, D. N., Bach, A. J. E., O'Brien, J. L., & Sainani, K. L. (2022). Calculating sample size for reliability studies. *PM&R*, *14*(8), 1018–1025. https://doi.org/10.1002/pmrj.12850

Chorba, R. S., Chorba, D. J., Bouillon, L. E., Overmyer, C. A., & Landis, J. A. (2010). Use of a functional movement screening tool to determine injury risk in female collegiate athletes. *North American Journal of Sports Physical Therapy: NAJSPT*, *5*(2), 47–54.

Comfort, P., Jones, Paul. A., McMahon, J. J., & Newton, R. (2015). Effect of Knee and Trunk Angle on Kinetic Variables During the Isometric Midthigh Pull: Test–Retest Reliability. *International Journal of Sports Physiology and Performance*, *10*(1), 58–63. https://doi.org/10.1123/ijspp.2014-0077

De la Motte, S. J., Gribbin, T. C., Lisman, P., Murphy, K., & Deuster, P. A. (2017). Systematic Review of the Association Between Physical Fitness and Musculoskeletal Injury Risk: Part 2—Muscular Endurance and Muscular Strength. *Journal of Strength and Conditioning Research*, *31*(11), 3218–3234. https://doi.org/10.1519/JSC.0000000000002174

Dello Iacono, A., Martone, D., & Padulo, J. (2016). Acute Effects of Drop-Jump Protocols on Explosive Performances of Elite Handball Players. *Journal of Strength and Conditioning Research*, *30*(11), 3122–3133. https://doi.org/10.1519/JSC.0000000000001393

Dennis, R. J., Finch, C. F., Elliott, B. C., & Farhart, P. J. (2008). The reliability of musculoskeletal screening tests used in cricket. *Physical Therapy in Sport*, *9*(1), 25–33. https://doi.org/10.1016/j.ptsp.2007.09.004

DeWolf, A., McPherson, A., Besong, K., Hiller, C., & Docherty, C. (2018). Quantitative Measures Utilized in Determining Pointe Readiness in Young Ballet Dancers. *Journal of Dance Medicine & Science*, *22*(4), 209–217. https://doi.org/10.12678/1089-313X.22.4.209

Evans, K., Refshauge, K. M., & Adams, R. (2007). Trunk muscle endurance tests: Reliability, and gender differences in athletes. *Journal of Science and Medicine in Sport*, *10*(6), 447–455. https://doi.org/10.1016/j.jsams.2006.09.003

Fernandez, M. R., Athens, J., Balsalobre-Fernandez, C., Kubo, M., & Hébert-Losier, K. (2023). Concurrent validity and reliability of a mobile iOS application used to assess calf raise test kinematics. *Musculoskeletal Science and Practice*, *63*, 102711. https://doi.org/10.1016/j.msksp.2022.102711

Gamboa, J. M., Roberts, L. A., Maring, J., & Fergus, A. (2008). Injury Patterns in Elite Preprofessional Ballet Dancers and the Utility of Screening Programs to Identify Risk Characteristics. *Journal of Orthopaedic & Sports Physical Therapy*, *38*(3), 126–136. https://doi.org/10.2519/jospt.2008.2390

Green, B., McClelland, J. A., Semciw, A. I., Schache, A. G., McCall, A., & Pizzari, T. (2022). The Assessment, Management and Prevention of Calf Muscle Strain Injuries: A Qualitative Study of the Practices and Perspectives of 20 Expert Sports Clinicians. *Sports Medicine - Open*, *8*(1), 10. https://doi.org/10.1186/s40798-021-00364-0

Haber, M., Golan, E., Azoulay, L., Kahn, S. R., & Shrier, I. (2004). Reliability of a device measuring triceps surae muscle fatigability. *British Journal of Sports Medicine*, *38*(2), 163–167. https://doi.org/10.1136/bjsm.2002.002899

Hébert-Losier, K., Newsham-West, R. J., Schneiders, A. G., & Sullivan, S. J. (2009). Raising the standards of the calf-raise test: A systematic review. *Journal of Science and Medicine in Sport*, *12*(6), 594–602. https://doi.org/10.1016/j.jsams.2008.12.628

Hébert-Losier, K., Ngawhika, T. M., Balsalobre-Fernandez, C., & O'Neill, S. (2023). Calf muscle abilities are related to sprint performance in male Rugby Union players. *Physical Therapy in Sport*, *64*, 117–122. https://doi.org/10.1016/j.ptsp.2023.09.001

Hébert-Losier, K., Wessman, C., Alricsson, M., & Svantesson, U. (2017). Updated reliability and normative values for the standing heel-rise test in healthy adults. *Physiotherapy*, *103*(4), 446–452. https://doi.org/10.1016/j.physio.2017.03.002

Hopkins, W. G. (2000). Measures of Reliability in Sports Medicine and Science: *Sports Medicine*, *30*(1), 1–15. https://doi.org/10.2165/00007256-200030010-00001

Howarth, D. J., Cohen, D. D., McLean, B. D., & Coutts, A. J. (2022). Establishing the Noise: Interday Ecological Reliability of Countermovement Jump Variables in Professional Rugby Union Players. *Journal of Strength and Conditioning Research*, *36*(11), 3159–3166. https://doi.org/10.1519/JSC.0000000000004037

Hunter, S. K., Duchateau, J., & Enoka, R. M. (2004). Muscle Fatigue and the Mechanisms of Task Failure: *Exercise and Sport Sciences Reviews*, *32*(2), 44–49. https://doi.org/10.1097/00003677-200404000-00002

Koo, T. K., & Li, M. Y. (2016). A Guideline of Selecting and Reporting Intraclass Correlation Coefficients for Reliability Research. *Journal of Chiropractic Medicine*, *15*(2), 155–163. https://doi.org/10.1016/j.jcm.2016.02.012

Liederbach, M. (1997). Screening for Functional Capacity in Dancers Designing Standardized, Dance-Specific Injury Prevention Screening Tools. *Journal of Dance Medicine & Science*, *1*(3), 93–106. https://doi.org/10.1177/1089313X9700100304

Lunsford, B. R., & Perry, J. (1995). The Standing Heel-Rise Test for Ankle Plantar Flexion: Criterion for Normal. *Physical Therapy*, *75*(8), 694–698. https://doi.org/10.1093/ptj/75.8.694

MacSweeney, N. D. H., Shaw, J. W., Simkin, G. P., Pedlar, C. R., Price, P. D. B., Mahaffey, R., & Cohen, D. D. (2024). Jumping Asymmetries and Risk of Injuries in Preprofessional Ballet. *The American Journal of Sports Medicine*, *52*(2), 492–502. https://doi.org/10.1177/03635465231218258

Mattiussi, A. M., Shaw, J. W., Cohen, D. D., Price, P., Brown, D. D., Pedlar, C. R., & Tallent, J. (2022). Reliability, variability, and minimal detectable change of bilateral and unilateral lower extremity isometric force tests. *The Journal of Sport and Exercise Science*, *6*(3). https://doi.org/10.36905/jses.2022.03.05

McKeown, I., Taylor-McKeown, K., Woods, C., & Ball, N. (2014). Athletic ability assessment: a movement assessment protocol for athletes. *International Journal of Sports Physical Therapy*, *9*(7), 862–873.

Möck, S., Hartmann, R., & Wirth, K. (2023). Vertical jumping performance relates to the one-repetition maximum in the standing calf raise and in the squat. *German Journal of Exercise and Sport Research*, *53*(2), 139–147. https://doi.org/10.1007/s12662-022-00852-9

Möller, M., Lind, K., Styf, J., & Karlsson, J. (2005). The reliability of isokinetic testing of the ankle joint and a heel-raise test for endurance. *Knee Surgery, Sports Traumatology, Arthroscopy*, *13*(1), 60–71. https://doi.org/10.1007/s00167-003-0441-0

Nunes, G. S., Tessarin, B. M., Scattone Silva, R., & Serrão, F. V. (2019). Relationship between the architecture and function of ankle plantar flexors with Achilles tendon morphology in ballet dancers. *Human Movement Science*, *67*, 102494. https://doi.org/10.1016/j.humov.2019.102494

Pires, I. M., Ponciano, V., Garcia, N. M., & Zdravevski, E. (2020). Analysis of the Results of Heel-Rise Test with Sensors: A Systematic Review. *Electronics*, *9*(7), 1154. https://doi.org/10.3390/electronics9071154

Rogers, D. K., McKeown, I., Parfitt, G., Burgess, D., & Eston, R. G. (2019). Inter- and Intra-rater Reliability of the Athletic Ability Assessment in Subelite Australian Rules Football Players. *Journal of Strength and Conditioning Research*, *33*(1), 125–138. https://doi.org/10.1519/JSC.0000000000002175

Ross, M. D., & Fontenot, E. G. (2000). Test–Retest Reliability of the Standing Heel-Rise Test. *Journal of Sport Rehabilitation*, *9*(2), 117–123. https://doi.org/10.1123/jsr.9.2.117

Rowley, K., Jarvis, D. N., Kurihara, T., Chang, Y.-J., Fietzer, A. L., & Kulig, K. (2015). Toe Flexor Strength, Flexibility and Function and Flexor Hallucis Longus Tendon Morphology in Dancers and Non-Dancers. *Medical Problems of Performing Artists*, *30*(3), 152–156. https://doi.org/10.21091/mppa.2015.3029

Sara, L. K., Gutsch, S. B., & Hunter, S. K. (2021). The single-leg heel raise does not predict maximal plantar flexion strength in healthy males and females. *PLOS ONE*, *16*(8), e0253276. https://doi.org/10.1371/journal.pone.0253276

Schrefl, A., Erlacher, D., & Schärli, A. (2023). The Single Leg Heel Rise Test—A Helpful Tool for Dance Science?: A Systematic Review. *Medical Problems of Performing Artists*, *38*(4), 234–248. https://doi.org/10.21091/mppa.2023.4028

Shechtman, O. (2013). The Coefficient of Variation as an Index of Measurement Reliability. In S. A. R. Doi & G. M. Williams (Eds.), *Methods of Clinical Epidemiology* (pp. 39–49). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-37131-8_4

Sman, A. D., Hiller, C. E., Imer, A., Ocsing, A., Burns, J., & Refshauge, K. M. (2014). Design and Reliability of a Novel Heel Rise Test Measuring Device for Plantarflexion Endurance. *BioMed Research International*, *2014*, 1–7. https://doi.org/10.1155/2014/391646

Thomas, K. S. (2003). Functional Élevé Performance as it Applies to Heel-Rises in Performance-Level Collegiate Dancers. *Journal of Dance Medicine & Science*, *7*(4), 115–120. https://doi.org/10.1177/1089313X0300700402

van Melick, N., Meddeler, B. M., Hoogeboom, T. J., Nijhuis-van Der Sanden, M. W. G., & van Cingel, R. E. H. (2017). How to determine leg dominance: The agreement between self-reported and observed performance in healthy adults. *PLOS ONE*, *12*(12), e0189876. https://doi.org/10.1371/journal.pone.0189876

Yeung, S. S., Suen, A. M. Y., & Yeung, E. W. (2009). A prospective cohort study of hamstring injuries in competitive sprinters: preseason muscle imbalance as a possible risk factor. *British Journal of Sports Medicine*, *43*(8), 589–594. https://doi.org/10.1136/bjsm.2008.056283

Zellers, J. A., Van Ostrand, K., & Silbernagel, K. G. (2017). Calf Endurance and Achilles Tendon Structure in Classical Ballet Dancers. *Journal of Dance Medicine & Science*, *21*(2), 64–69. https://doi.org/10.12678/1089-313X.21.2.64

Acknowledgments

We want to thank Fritz Zaunrieth for designing and building the laser pointer device for our measurements.

Funding

The authors have no funding or support to report.

Competing interests

The authors have declared that no competing interests exist.

Data availability statement

All relevant data are within the paper.