Article

Intra-Rater and Inter-Rater Reliability of Hand-Held Dynamometry for Shoulder Strength Assessment in Circus Arts Students

Charlotte Ganderton, PhD,[1] Bonnie Kerr, M.Physio,[2] Matthew King, PhD,[3] Ross Lenssen, BHSc(Physiother),[2] Sarah Warby, PhD,[1,2] David Munro, PhD,[4] Lyn Watson, Prof.Doc,[2] Simon Balster, BPhty(Hons),[2] Jia Han, PhD,[1,5,6] and Oren Tirosh, PhD[5]

The purpose of this study was to establish the intra-rater and inter-rater reliability of isometric shoulder strength assessment using a hand-held dynamometer (HHD) in functional joint positions in student circus artists with symptomatic atraumatic shoulder instability. **METHODS:** Over two testing sessions, two experienced physiotherapists assessed the shoulder strength of 24 student circus artists with clinically diagnosed atraumatic shoulder instability. Both the symptomatic and asymptomatic shoulder was assessed using a HHD in 10 functional positions. Intra-class correlation coefficients (ICCs) were calculated to determine the reliability of strength measurements. **RESULTS:** All examined positions showed moderate-high intra-rater and inter-rater reliability. External rotation at 0° and internal rotation in horizontal flexion at 45° revealed the most reliable results, and the shrug position the least reliable. Inter-rater and intra-rater reliability was high and demonstrated similar results in symptomatic and asymptomatic shoulders by both raters. **DISCUSSION:** This study demonstrated clinical applicability in reliably measuring functional strength in symptomatic atraumatic instability or asymptomatic shoulders when assessed by experienced therapists using an HHD. *Med Probl Perform Art* 2021;36(2):88–102.

CIRCUS is a physically demanding performing arts discipline that requires high levels of stability, strength, and repetitive force attenuation in positions of large shoulder joint range of motion (ROM).[1] The prevalence of shoulder injuries in circus artists is high,[2,3] representing a yearly prevalence of 27.7% at the University of Arts in the Netherlands[2] and 12% at the National Institute of Circus Arts (NICA) in Australia.[1] Testing shoulder strength in symptomatic[4,5] and asymptomatic[5–13] shoulders has been investigated in a variety of sports, predominantly in positions of low shoulder joint range of motion,[14,15] but there remains a dearth of research in the performing arts and in mid to end of range joint positions.

A comprehensive shoulder assessment assists in attaining a diagnosis, directs rehabilitation and activity modification, and guides return to activity post-injury.[6] Additionally, reliable strength assessments provide valuable clinical insight for targeted injury prevention strategies and to monitor the efficacy of interventions.[6] Strength can be assessed using an isokinetic dynamometer (ID), externally fixed dynamometer (EFD), or hand-held dynamometer (HHD). Isokinetic dynamometry is reported as the gold standard of strength assessment, with excellent reliability and validity.[12,16–18] It allows maximal force generation without relying on the matched force of the therapist; though, its utility is limited in circus arts due to its large size, lack of portability, and high cost. Therefore, there is a need for a less cumbersome, but reliable, assessment technique.

HHD may provide a viable method of shoulder strength assessment for circus artists. Its portable and light-weight nature provides benefits in both times of high touring demand and the ability to test a variety of movements and positions. When compared to ID, objective strength measurements obtained via HHD have demonstrated high intra- and inter-rater reliability across a variety of cohorts, albeit only in relatively neutral joint positions (e.g., shoulder external rotation in neutral internal and external rotation[5–13, 19]). Similar to swimming[20,21] and gymnastics,[15,22] where there is a high prevalence of joint hypermobility and atraumatic shoulder instability, the implications of instability on shoulder strength assessment is unknown in circus performers. Thus, there is a need for research that investigates shoulder strength in mid to end-range positions.

From the [1]School of Health Sciences, Swinburne University of Technology, Hawthorn, Victoria, Australia; [2]Melbourne Shoulder Group, Prahran, Victoria, Australia; [3]La Trobe Sports and Exercise Medicine Research Centre, La Trobe University, Bundoora, Victoria, Australia; [4]National Institute of Circus Arts, Prahran, Victoria, Australia; [5]Shanghai University of Sport, Shanghai, China; and [6]University of Canberra, Australia.

The authors declare no funding or conflicts of interest related to this study.

Address for correspondence: Dr. Charlotte Ganderton, Swinburne University of Technology, John Street, Hawthorn, VIC 3122, Australia. Tel +61 3 9214 3539. cganderton@swin.edu.au.

https://doi.org/10.21091/mppa.2021.2014

The functional relevance of strength testing positions is a crucial consideration in assessing individuals exposed to repetitive application of high joint forces in large ranges of shoulder movement. However, the lack of research in this area limits understanding of appropriate assessment techniques. It inhibits the ability to make meaningful clinical recommendations regarding the reliability of HHD for measuring shoulder strength. Therefore, the primary aim of this study was to determine the intra-rater and inter-rater reliability of functional isometric shoulder strength assessment using HHD in student circus artists with unilateral symptomatic atraumatic instability.

METHODS

A convenience sample of 24 adult NICA students with clinically diagnosed atraumatic shoulder instability were included in the study. One independent therapist diagnosed shoulder instability in all students. Atraumatic shoulder instability was defined as having discomfort, pain, apprehension, and guarding with movement of the glenohumeral joint in one or more directions, for at least one positive clinical test finding, in the absence of a significant traumatic event.[23,24] Instability was determined using the following clinical assessments: Sulcus sign,[25–28] anterior and posterior draw test (10–30° abduction),[25,29] anterior and posterior drawer test abducted (80–120° abduction),[25,29] the anterior[26,30,31] and posterior[27,32] apprehension test or active movement dysfunction. Active movement dysfunction was assessed using visual observation and palpation during active movement, involving displacement or loss of centering of the humeral head during abduction, flexion, horizontal flexion with or without internal rotation, horizontal extension with external rotation[23][Appendix A]. These assessments have been shown to be valid and reliable and used in current literature.[23,33] The Beighton score[34] for global joint hypermobility was also recorded for each participant.

Participants were excluded if they had a history of significant trauma (e.g. glenohumeral dislocation requiring relocation, surgery); non-correctable volitional instability (deliberate dislocation/relocation); extreme anxiety (medical history); neurological motor deficit (medical history); instability due to upper motor neuron or lower motor neuron lesion (medical history); Ehler-Danlos syndrome/Marfan's syndrome (medical history); vascular thoracic outlet syndrome (clinical signs or symptoms); or a cervical referred signs or symptoms (using Spurling's[35] test). All participants provided written informed consent, and ethical approval was granted by the Swinburne University Human Research Ethics Committee (20202674-4451).

Testing Procedure

A JTECH Commander Echo Digital Dynamometer (JTECH Medical, Utah, USA) with a concave testing surface was used for all testing procedures. Two male physiotherapists, each with over 10 years of clinical experience working exclusively in shoulder assessment and rehabilitation, performed strength assessments. They both had previous experience using HHD in patients with atraumatic shoulder instability. Therapists were not blinded to the participants' symptomatic side during testing; however, they were blinded to each other's results. Anthropometrics were obtained, including height, measured using a measuring tape, and weight, measured using scales.

Participants were involved in two testing sessions, each completed at the same time of the day, 4 days apart. Therapist order and order of testing positions remained consistent over the two testing sessions. Prior to testing, each therapist placed a mark 5 cm proximal to the wrist joint line on the volar and palmer aspect of the forearm. This acted as a reference point for dynamometer placement (Table 1). A 'make test' protocol was decided a priori to mitigate any potential risks associated with the testing procedure in a clinically unstable shoulder population. Testing was performed on a gym floor without shoes. Participants completed three maximal test efforts with a 10-second rest between repetitions and a 30-second rest between test positions. Standardised testing positions (see Table 1, Appendix B) were used for each of the 10 tests: (1) ER at 0°, (2) IR at 0°, (3) ER abducted at 90°, (4) IR abducted at 90°, (5) abduction at 45°, (6) ER in horizontal flexion 45°, (7) IR in horizontal flexion 45°, (8) flexion at 90°, (9) extension at 90°, (10) elevation (shrug). The positions tested were developed through consultation with expert shoulder clinicians, including the authors of a recently published randomised controlled trial evaluating the efficacy of two rehabilitation programs for multidirectional instability of the glenohumeral joint.[36]

Each testing position was explained and demonstrated by the therapists, and the participant was allowed one practice (sub-maximal contraction) prior to recorded measurements. Standardised encouragement was given to the participants via a recording of a monotonous voice with the phrase "go ahead . . . push, push, push, push, push, relax." This was used to ensure consistent encouragement and time spent at each position.[37] All participants were tested on their symptomatic and asymptomatic sides in a consistent order of left shoulder followed by right shoulder. The therapist position was not standardised further to the information provided in Table 1. This allowed the therapist to stand in a way that optimised a stable base of support. Where possible, therapists palpated for humeral head translation but did not forcefully stabilise the joint in any way. The first testing session was used to measure inter-rater reliability, and a 10-minute rest was provided between therapists. Participants had no exposure to shoulder strength testing at NICA prior to the undertaking this study. No participants reported pain during testing.

Data Analysis

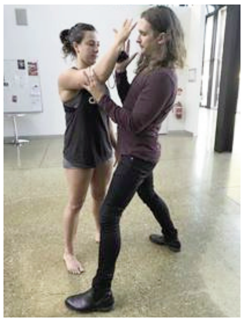
The best of three repetitions (absolute peak force) and mean of the three repetitions (average peak force) for

TABLE 1. Hand-held Dynamometer Testing Procedure

ER@0° 	Standing: feet parallel hip-width apart, torso neutral	Volar aspect of the distal forearm,* force applied perpendicular to the testing limb	Arm neutral by side, elbow bent at 90° flexion, wrist in anatomical neutral	To the affected side; one hand on the HHD and one hand on the joint line
IR@0° 	Standing: feet parallel hip-width apart, torso neutral	Palmar aspect of the distal forearm,* force applied perpendicular to the testing limb	Arm neutral by side, elbow bent at 90° flexion, wrist in anatomical neutral	To the affected side; one hand on the HHD and one hand on the joint line
ER@90° 	Standing: split stance, opposite leg forward, torso neutral	Volar aspect of the distal forearm,* force applied perpendicular to the testing limb	Arm at 90° abd, ER to 90°, 90° elbow flexion, wrist in anatomical neutral	Behind; one hand on the HHD and one hand on the joint line
IR@90° 	Standing: split stance, opposite leg forward, torso neutral	Palmar aspect of the distal forearm,* force applied perpendicular to the testing limb	Arm at 90° abd, ER to 90°, 90° elbow flexion, wrist in anatomical neutral	In front; one hand on the HHD and one hand on the joint line
Abd@45° 	Standing: feet parallel hip-width apart, torso neutral	Lateral aspect of the distal arm, force applied perpendicular to the testing limb	Arm at 45° abduction, neutral rotation, 90° elbow flexion, wrist in anatomical neutral	To the affected side; one hand on the HHD and one hand on the joint line

(continued on next page)

TABLE 1. Hand-held Dynamometer Testing Procedure (*continued*)

Test Position	Participant Position	Dynamometer Position	Arm Position	Therapist Position
ERinHF@45° 	Standing: feet parallel hip-width apart, torso neutral	Volar aspect of the distal forearm,* force applied perpendicular to the testing limb	Arm at 90° flexion, IR to 45°, 90° elbow flexion, wrist in anatomical neutral	To the affected side; one hand on the HHD and one hand stabilising the elbow
IRinHF@45° 	Standing: feet parallel hip-width apart, torso neutral	Palmar aspect of the distal forearm,* force applied perpendicular to the testing limb	Arm at 90° flexion, IR to 45°, 90°elbow flexion, wrist in anatomical neutral	To the unaffected side; one hand on the HHD and one hand stabilising the elbow
Ext@90° 	Prone	Posterior aspect of the distal arm, force applied perpendicular to the testing limb	Arm at 90° abd, neutral rotation, 90° elbow flexion, wrist in anatomical neutral	To the affected side; one hand on the HHD, one hand stabilising the torso
Flex@90° 	Kneeling: split stance, opposite leg forward, torso neutral	On the fist, force applied directly to the testing limb	Arm at 90° flexion, neutral rotation, 90° elbow flexion, wrist in neutral, hand in a flexed fist	To the affected side; stabilising the HHD with two hands
Shrug 	Standing: feet parallel hip-width apart, torso neutral	Top of scapular, force applied directly to the testing limb, stabilised with external fixation	Arm held at 30° abd, arm and wrist in anatomical neutral, palm facing forwards	To the affected side; stabilising the HHD with one hand

(footnotes on next page)

Notes to Table 1

*The dynamometer was placed 5 cm proximal to the wrist joint.

Note: Images displayed are to demonstrate testing positions only. Data were collected in the order of appearance in the table. HHD: hand-held dynamometer; ER: external rotation; IR: internal rotation; Abd: abduction; Ext: Extension.

ER@0°: external rotation at 0° of shoulder abduction; IR@0°: internal rotation at 0° of shoulder abduction; ER@90°: external rotation at 90° of shoulder abduction; IR@90°: internal rotation at 90° of shoulder abduction; Abd@45°: abduction in 45° of shoulder abduction; ERinHF@45°: external rotation in 45° of horizontal flexion; IRinHF@45°: internal rotation in 45° of horizontal flexion, Ext@90°: extension in 90° of shoulder abduction; Flex@90°: flexion at 90° of shoulder abduction.

asymptomatic and symptomatic shoulders were recorded in Newtons (N) for each testing position. Data distribution was assessed using a Shapiro-Wilk test, Q-Q plots, and visual inspection of boxplots, and summarised using means and standard deviations (SD) or medians and interquartile ranges [IQR] as appropriate. Paired *t*-tests were used to analyse systematic differences between day 0 and day 4 testing sessions. Intra-rater reliability of each tester were assessed across the two testing sessions using the average of the three repetitions taken with an intraclass correlation coefficient ($ICC_{3,1}$, two-way model, consistency definition) and 95% confidence interval (95%CI).[38,39]

Inter-rater reliability was analysed by using ICC ($ICC_{2,1}$ calculated two-way model, absolute definition) highlighting differences between raters for the symptomatic and asymptomatic shoulders.[38] Clinical significance of the ICC was interpreted according to the methods of Myers and Blesh,[40] where 0.60–0.69 was considered poor, 0.70–0.79 fair, 0.80–0.89 good, and 0.90 and above high.

The standard error of measurement (SEM) was calculated by using the SD of between-group differences (SD × √(1–ICC)), where SD was the SD of all participant scores.[39] The SEM values were then used to calculate the minimum detectable change (MDC) and smallest detectable change at the 95% CI (SEM × 1.96 × √2) for all test positions.[39] The MDC was defined as the smallest statistically significant change that can be attributed to factors beyond measurement error, to ensure an objective change in strength.[41] If the difference in force between day 0 and day 4 was greater than the MDC, it was attributed to a change in strength.[41] A *p*-value of 0.05 and SEM <10% was chosen to indicate statistical significance (with a 95%CI). To aid in clinical interpretability, SEMs were converted to percentages of the first testing session for peak and mean force (Appendix C). All statistical analyses were conducted using Statistical Package for the Social Sciences (SPSS) version 26.0 (IBM SPSS, Armonk, NY, USA).

RESULTS

Participants included 8 males [mean (SD) age 23.39 yrs (SD 2.59); height 1.7 m (0.06) body mass 72.65 kg (11.18); median Beighton score 4 (IQR 2.5)] and 16 females [age 24.45 yrs (3.03); height 1.62 m (0.04); body mass 59.81 kg (6.67); median Beighton score 5.5 (IQR 5.5)]. All participants completed both testing sessions, and no adverse events were reported.

Intra-Rater Reliability

Intra-rater reliability results are outlined in Table 2. All examined testing positions displayed moderate-high intra-rater reliability results. The $ICC_{2,1}$ (95%CI) values showed good to high intra-rater reliability for both rater 1 (R1) and rater 2 (R2), with $ICC_{3,1}$ values ranging from 0.84–0.96 for absolute peak force and 0.83–0.94 for the average peak force across testing positions. Testing positions resulting in force below <110N (ER@0° and ER90°) had better intra-rater reliability using absolute peak force [ER@0°, 0.90 (0.76–0.96) and ER90°, 0.91 (0.78–0.96)] compared with the shrug position. The shrug test had the highest force (>300N) for symptomatic and asymptomatic shoulders of all positions in both testing sessions and had good to high intra-rater reliability ($ICC_{2,1}$ 0.85–94), though large 95%CI (e.g., 0.58–0.92).

The SEM values ranged from 1.26 to 12.52 for all positions except the shrug, which had significantly higher SEM values. When using absolute peak force, measurement variation (%SEM) was acceptable (<10%) for all tests except for IR@90° (12%) for the symptomatic shoulder and IR@90° (12%), shrug (13%), and ER@90° (10%) for the asymptomatic shoulder (Appendix C). Similarly, %SEM was acceptable for all tests when using the average peak force for all testing positions except IR@90° (11%) and shrug (12%) for the symptomatic shoulder and IR@90° (12%) and shrug (12%) for the asymptomatic shoulder (Appendix C).

Rater 2 showed consistently higher $ICC_{3,1}$ than R1, using average peak force, as evidenced during ER@0° (R2 93.36N, R1 85.78N), IR@90° (R2 92.31N, R1 86.88N) and shrug (R2 356.42N, R1 311.02N). There were some discrepancies in the direction of change between testing sessions: R2 showed an increase in absolute peak force during ER@90° [18.76 (11.38)] and ERinHF@45° [14.12 (8.97)] for the affected arm, compared to R1 that showed decrease [ER@90°, 7.37 (10.26)]. Similar findings were identified using the average peak force.

Inter-Rater Reliability

Inter-rater reliability results are outlined in Table 3. The $ICC_{2,1}$ (95%CI) values reveal all strength tests performed have clinically acceptable agreement (<0.7) based on the criteria by Meyers and Blesh.[40] High inter-rater reliability[40] was observed using absolute peak force, in positions of ER@0° [0.91 (0.77–0.96)], IR@0° [0.94 (0.84–0.98)] and IRinHF@45° [0.93 (0.84–0.97)] for both symptomatic and

TABLE 2. Inter-Rater Reliability Results

	Rater	Arm	Measurement	Difference (day 0–day 4)*	Paired *t*-test	$ICC_{2,1}$ (95%CI) (inter-rater reliability)	SEM (95%CI) [in Newtons]	MDC (95%CI) [in Newtons]
ER@0°	1	Affected	Best of 3 reps.	−3.86 (14.45)	0.214	0.90 (0.76, 0.96)	4.57 (2.89, 7.08)	12.67 (8.01, 19.62)
			(M) 3 best reps.	−1.83 (11.77)	0.463	0.92 (0.81, 0.97)	3.33 (2.04, 5.13)	9.23 (5.65, 14.22)
		Unaffected	Best of 3 reps.	2.13 (11.77)	0.396	0.94 (0.85, 0.97)	2.88 (2.04, 4.56)	7.99 (5.65, 12.64)
			(M) 3 best reps.	2.65 (11.57)	0.285	0.93 (0.83, 0.97)	3.06 (2.00, 4.77)	8.49 (5.55, 13.22)
	2	Affected	Best of 3 reps.	−9.09 (11.87)	0.001	0.92 (0.83, 0.97)	3.36 (2.06, 4.89)	9.31 (5.70, 13.57)
			(M) 3 best reps.	−7.41 (10.57)	0.003	0.94 (0.85, 0.97)	2.59 (1.83, 4.09)	7.18 (5.07, 11.35)
		Unaffected	Best of 3 reps.	−7.16 (8.03)	<0.001	0.97 (0.92, 0.99)	1.39 (0.8, 2.27)	3.86 (2.23, 6.30)
			(M) 3 best reps.	−6.51 (7.30)	<0.001	0.97 (0.93, 0.99)	1.26 (0.73, 1.93)	3.50 (2.02, 5.35)
IR@0°	1	Affected	Best of 3 reps.	11.79 (29.78)	0.071	0.89 (0.75, 0.95)	9.88 (6.66, 14.89)	27.38 (18.46, 41.27)
			(M) 3 best reps.	11.48 (24.87)	0.038	0.92 (0.81, 0.97)	7.03 (4.31, 10.84)	19.50 (11.94, 30.05)
		Unaffected	Best of 3 reps.	10.55 (24.91)	0.055	0.92 (0.80, 0.96)	7.05 (4.98, 11.14)	19.53 (13.81, 30.88)
			(M) 3 best reps.	9.59 (21.31)	0.042	0.93 (0.84, 0.97)	5.64 (3.69, 8.52)	15.63 (10.23, 23.63)
	2	Affected	Best of 3 reps.	−9.87 (16.20)	0.008	0.96 (0.90, 0.98)	3.24 (2.29, 5.12)	8.98 (6.35, 14.20)
			(M) 3 best reps.	−9.99 (18.36)	0.016	0.94 (0.86, 0.98)	4.50 (2.60, 6.87)	12.47 (7.20, 19.04)
		Unaffected	Best of 3 reps.	−11.60 (16.07)	0.003	0.96 (0.91, 0.99)	3.21 (1.61to 4.82)	8.91 (4.45, 13.36)
			(M) 3 best reps.	−12.24 (14.77)	0.001	0.96 (0.92, 0.99)	2.95 (2.95, 4.18)	8.19 (8.19, 11.58)
ER@90°	1	Affected	Best of 3 reps.	7.37 (10.26)	0.002	0.91 (0.78, 0.96)	3.08 (2.05, 4.81)	8.53 (5.69, 13.34)
			(M) 3 best reps.	8.20 (10.26)	0.001	0.89 (0.74, 0.95)	3.40 (2.29, 5.23)	9.43 (6.36, 14.50)
		Unaffected	Best of 3 reps.	7.73 (15.49)	0.026	0.82 (0.59, 0.93)	6.57 (4.10, 9.92)	18.22 (11.36, 27.49)
			(M) 3 best reps.	6.44 (11.71)	0.015	0.87 (0.70, 0.95)	4.22 (2.62, 6.41)	11.70 (7.26, 17.79)
	2	Affected	Best of 3 reps.	−18.76 (11.38)	<0.001	0.89 (0.75, 0.96)	3.77 (2.28, 5.69)	10.46 (6.31, 15.77)
			(M) 3 best reps.	−17.72 (9.71)	<0.001	0.91 (0.79, 0.96)	2.91 (1.94, 4.45)	8.07 (5.38, 12.33)
		Unaffected	Best of 3 reps.	−18.76 (10.64)	<0.001	0.88 (0.71, 0.95)	3.69 (2.38, 5.73)	10.22 (6.59, 15.88)
			(M) 3 best reps.	−18.36 (10.06)	<0.001	0.89 (0.74, 0.95)	3.34 (2.25, 5.13)	9.25 (6.24, 14.22)
IR@90°	1	Affected	Best of 3 reps.	3.56 (28.21)	0.551	0.84 (0.61, 0.93)	11.28 (7.46, 17.62)	31.28 (20.69, 48.83)
			(M) 3 best reps.	1.84 (24.35)	0.720	0.84 (0.62, 0.93)	9.74 (6.44, 15.01)	27.00 (17.86, 41.61)
		Unaffected	Best of 3 reps.	7.36 (32.32)	0.287	0.85 (0.64, 0.94)	12.52 (7.92, 19.39)	34.70 (21.94, 53.75)
			(M) 3 best reps.	8.60 (29.26)	0.173	0.83 (0.59, 0.93)	12.06 (7.74, 18.74)	33.44 (21.46, 51.93)
	2	Affected	Best of 3 reps.	−24.75 (17.52)	<0.001	0.94 (0.85, 0.97)	4.29 (3.03, 4.29)	11.90 (8.41, 11.90)
			(M) 3 best reps.	−22.33 (16.98)	<0.001	0.93 (0.84, 0.97)	4.49 (2.94, 4.49)	12.45 (8.15, 12.45)
		Unaffected	Best of 3 reps.	−25.33 (16.87)	<0.001	0.93 (0.84, 0.97)	4.46 (2.92, 6.75)	12.37 (8.10, 18.70)
			(M) 3 best reps.	−22.88 (15.35)	<0.001	0.94 (0.85, 0.97)	3.76 (3.76, 5.95)	10.42 (10.42, 16.48)
Shrug	1	Affected	Best of 3 reps.	−67.91 (88.87)	0.001	0.85 (0.64, 0.94)	34.42 (21.77, 53.32)	95.41 (60.34, 147.80)
			(M) 3 best reps.	−65.51 (86.70)	0.002	0.83 (0.60, 0.93)	35.75 (22.94, 54.83)	99.09 (63.58, 151.99)
		Unaffected	Best of 3 reps.	−53.06 (100.35)	0.019	0.82 (0.58, 0.92)	46.82 (31.21, 71.51)	129.77 (86.51, 198.23)
			(M) 3 best reps.	−57.23 (93.31)	0.008	0.83 (0.59, 0.93)	38.47 (24.69, 59.75)	106.64 (68.43, 165.61)
	2	Affected	Best of 3 reps.	−40.02 (59.65)	0.004	0.94 (0.87, 0.98)	14.61 (8.44, 21.51)	40.50 (23.38, 59.61)
			(M) 3 best reps.	−33.97 (55.22)	0.007	0.94 (0.87, 0.98)	13.53 (7.81, 19.91)	37.49 (21.65, 55.19)
		Unaffected	Best of 3 reps.	−18.27 (53.54)	0.116	0.96 (0.91, 0.98)	10.71 (7.57, 16.06)	29.68 (20.99, 44.52)
			(M) 3 best reps.	−23.99 (62.18)	0.073	0.94 (0.85, 0.97)	15.23 (10.77, 24.08)	42.22 (29.85, 66.75)
ERinHF@45°	1	Affected	Best of 3 reps.	4.64 (15.02)	0.153	0.86 (0.67, 0.94)	5.62 (3.68, 8.63)	15.58 (10.20, 23.92)
			(M) 3 best reps.	2.43 (11.15)	0.307	0.91 (0.79, 0.96)	3.35 (2.23, 5.11)	9.27 (6.18, 14.16)
		Unaffected	Best of 3 reps.	−2.67 (10.13)	0.219	0.90 (0.77, 0.96)	3.20 (2.03, 4.86)	8.88 (5.62, 13.47)
			(M) 3 best reps.	−0.93 (9.61)	0.648	0.88 (0.72, 0.95)	3.33 (2.15, 5.09)	9.23 (5.96, 14.10)
	2	Affected	Best of 3 reps.	−14.12 (8.97)	<0.001	0.94 (0.87, 0.98)	2.20 (1.27, 3.23)	6.09 (3.52, 8.96)
			(M) 3 best reps.	−12.77 (8.67)	<0.001	0.94 (0.86, 0.98)	2.12 (1.23, 3.24)	5.89 (3.40, 8.99)
		Unaffected	Best of 3 reps.	−12.96 (12.29)	<0.001	0.87 (0.70, 0.95)	4.43 (2.75, 6.73)	12.28 (7.62, 18.66)
			(M) 3 best reps.	−13.87 (10.63)	<0.001	0.89 (0.75, 0.95)	3.53 (2.38, 5.32)	9.77 (6.59, 14.73)

(continued on next page)

TABLE 2. Inter-Rater Reliability Results (*continued*)

	Rater	Arm	Measurement	Difference (day 0–day 4)*	Paired *t*-test	$ICC_{2,1}$ (95%CI) (inter-rater reliability)	SEM (95%CI) [in Newtons]	MDC (95%CI) [in Newtons]
IRinHF@45°	1	Affected	Best of 3 reps.	−7.10 (35.43)	0.347	0.91 (0.78, 0.96)	10.63 (7.09 to16.62)	29.46 (19.64, 46.06)
			(M) 3 best reps.	−4.72 (32.75)	0.497	0.91 (0.76, 0.96)	9.83 (6.55, 16.04)	27.23 (18.16, 44.47)
		Unaffected	Best of 3 reps.	−5.22 (35.98)	0.494	0.92 (0.82, 0.97)	10.11 (6.19, 15.17)	28.03 (17.16, 42.04)
			(M) 3 best reps.	−0.37 (31.13)	0.955	0.96 (0.90, 0.98)	6.23 (9.84, 4.40)	17.26 (12.20, 27.29)
	2	Affected	Best of 3 reps.	−15.08 (21.44)	0.003	0.96 (0.91, 0.98)	4.29 (3.03, 6.43)	11.89 (8.40, 17.83)
			(M) 3 best reps.	−18.90 (19.75)	<0.001	0.92 (0.82, 0.96)	5.59 (3.95, 8.38)	15.48 (10.95, 23.23)
		Unaffected	Best of 3 reps.	−19.14 (22.17)	<0.001	0.95 (0.89, 0.98)	4.96 (3.14, 7.35)	13.74 (8.69, 20.38)
			(M) 3 best reps.	−17.98 (19.83)	<0.001	0.96 (0.90, 0.98)	3.97 (2.80, 6.27)	10.99 (7.77, 17.38)

* For Difference day 1–4, data given as mean n (SD). For SEM and MDC, data given in Newtons. ER: external rotation; IR: internal rotation; HF: horizontal flexion; ICC: intra-class correlation coefficient; SEM: standard error of measurement; MDC: minimal detectable change; M: mean; reps.: repetitions. See Supplementary Appendix C for full data collection.

ER@0°: external rotation at 0° of shoulder abduction; IR@0°: internal rotation at 0° of shoulder abduction; ER@90°: external rotation at 90° of shoulder abduction; IR@90°: internal rotation at 90° of shoulder abduction; Abd@45°: abduction in 45° of shoulder abduction; ERinHF@45°: external rotation in 45° of horizontal flexion; IRinHR@45°: internal rotation in 45° of horizontal flexion, Ext@90°: extension in 90° of shoulder abduction; Flex@90°: flexion at 90° of shoulder abduction.

asymptomatic shoulders. Good intra-rater reliability scores[40] were observed in positions ER@90°, IR@90°, and ERinHF@45° for both the symptomatic and asymptomatic shoulders.

The shrug position had fair to good inter-rater reliability for the symptomatic shoulder and good inter-rater reliability for the asymptomatic shoulder. Narrow 95%CI and lower bound CI >0.70 existed for all test positions on the symptomatic shoulder, except for the shrug ($ICC_{2,1}$ 0.80, 95%CI, 0.44–0.92).

The study provides similar results in both inter-rater and intra-rater reliability between the symptomatic and asymptomatic shoulder. Analysing the $ICC_{3,1}$ of R1 with ER@0° and IR@0° demonstrated similar results between symptomatic (0.92/0.92) and asymptomatic (0.93/0.93) shoulder as did R2, symptomatic (0.94/0.94) and asymptomatic (0.97/0.96) using the average peak force. Inter-rater $ICC_{2,1}$ of R1 with ER@0° and IR@0° also demonstrate similarities between symptomatic (0.91/0.96) and asymptomatic (0.94/0.91) using the average peak force.

DISCUSSION

The results of this study show good-high reliability for the use of HHD to assess shoulder strength for all positions tested when repeated measures are taken by the same clinician. This presents important clinically applicable results for clinicians that may be generalisable to symptomatic circus or other performing artists or athletes presenting with atraumatic instability when performing assessments or monitoring rehabilitation progress.

Testing positions in lower ranges of shoulder motion tended to have greater inter-rater and intra-rater reliability than those in higher ranges. This may be due to increased humeral head translation in greater ranges of shoulder elevation and a biomechanical disadvantage in the therapist's ability to stabilise in these positions.

The shrug testing position had a force >300N and had lower intra-rater reliability, higher variability in the 95%CI, high SEM and large 95%CI compared to all other testing positions. It has been proposed that the limit of HHD is 300N,[42] due to the difficulty of stabilising higher levels of force. Previous research by Thorburg et al.[43] showed comparable results when evaluating HHD on hip strength testing, where positions of force >300N had lower intra-rater reliability. Furthermore, the HHD was fixed (Table 1), which enables the participant to compensate, reducing reliability.

Fair to excellent inter-rater reliability was shown in the results of this study. Inter-rater reliability differed between R1 and R2, with R2 showing consistently higher $ICC_{3,1}$ than R1. Furthermore, discrepancies existed in the direction of change evident in some test positions (ER@90°, ERinHR@45°). Differences between raters may be attributed to therapist error, as identified differences are predominantly lower than the MDC. Tester variance could be caused by tester stabilisation, anthropometrics, or subtle differences in HDD placement across sessions.

The reliability of strength testing may be altered depending on if assessing a symptomatic or asymptomatic shoulder. Joint hypermobility syndrome was a strong characteristic across the included cohort, with 58.3% of participants satisfying the Beighton criteria.[34] This is comparable to professional ballet dancers, where hypermobility is common in males and females.[14] Evidence suggests a relationship between generalised joint hypermobility and atraumatic instability, which is thought to be due to poor afferent input, altered neuromuscular control and proprioception.[44] Muscular stabilisation ability, power capability and hypermobility in large ranges of movement are required in the circus arts, which is thought to increase instability and shoulder injury.[45] The reliability of force output in shoulder instability may be due to a less variable motor activation strategy.[46] This study showed high reliability in symptomatic and asymptomatic shoulders across

TABLE 3. Inter-Rater Reliability Results

	Arm	Measurement	$ICC_{2,1}$ (95%CI) (inter-rater reliability)	SEM (95%CI) [in Newtons]	MDC (95%CI) [in Newtons]
ER@0°	Affected	Best of 3 reps.	0.91 (0.77, 0.96)	7.19 (4.79, 11.49)	19.93 (13.29, 31.86)
		(M) 3 best reps.	0.91 (0.75, 0.96)	6.92 (4.61, 11.53)	19.18 (12.79, 31.97)
	Unaffected	Best of 3 reps.	0.94 (0.85, 0.97)	7.03 (4.97, 11.12)	19.49 (13.78, 30.81)
		(M) 3 best reps.	0.94 (0.80, 0.98)	6.57 (3.29, 11.99)	18.21 (10.51, 33.25)
IR@0°	Affected	Best of 3 reps.	0.94 (0.84, 0.98)	12.20 (7.05, 19.93)	33.82 (19.53, 55.23)
		(M) 3 best reps.	0.96 (0.90, 0.98)	9.11 (6.44, 14.40)	25.24 (17.85, 39.91)
	Unaffected	Best of 3 reps.	0.92 (0.80, 0.96)	14.69 (10.39, 23.22)	40.71 (28.79, 64.37)
		(M) 3 best reps.	0.91 (0.78, 0.96)	14.42 (9.62, 22.55)	39.98 (26.65, 62.51)
ER@90°	Affected	Best of 3 reps.	0.87 (0.71, 0.95)	7.05 (4.37, 10.53)	19.54 (12.12, 29.19)
		(M) 3 best reps.	0.89 (0.75, 0.95)	6.07 (4.09, 9.15)	16.82 (11.34, 25.36)
	Unaffected	Best of 3 reps.	0.82 (0.58, 0.92)	8.50 (5.67, 12.99)	23.67 (15.71, 36.00)
		(M) 3 best reps.	0.86 (0.67, 0.94)	6.66 (4.36, 10.23)	18.47 (12.09, 28.35)
IR@90°	Affected	Best of 3 reps.	0.87 (0.70, 0.95)	13.75 (8.30, 20.32)	37.08 (23.00, 56.33)
		(M) 3 best reps.	0.83 (0.76, 0.95)	13.75 (7.46, 16.33)	38.10 (20.67, 45.28)
	Unaffected	Best of 3 reps.	0.86 (0.65, 0.94)	15.05 (9.85, 23.79)	41.70 (27.30, 65.94)
		(M) 3 best reps.	0.89 (0.74, 0.95)	12.14 (8.18, 18.66)	33.65 (22.68, 51.73)
Shrug	Affected	Best of 3 reps.	0.80 (0.44, 0.92)	49.43 (31.26, 82.72)	137.01 (86.66, 229.28)
		(M) 3 best reps.	0.79 (0.30, 0.92)	47.51 (29.33, 86.75)	131.70 (81.29, 240.46)
	Unaffected	Best of 3 reps.	0.81 (0.43, 0.93)	55.68 (33.80, 96.44)	154.34 (93.68, 267.32)
		(M) 3 best reps.	0.82 (0.37, 0.93)	49.07 (30.60, 91.80)	135.00 (84.81, 254.44)
ERinHF@45°	Affected	Best of 3 reps.	0.88 (0.72, 0.95)	7.81 (4.63, 10.97)	19.90 (12.85, 30.40)
		(M) 3 best reps.	0.87 (0.70, 0.95)	6.82 (4.24, 10.38)	18.93 (11.74, 28.76)
	Unaffected	Best of 3 reps.	0.82 (0.43, 0.93)	7.91 (4.93, 14.08)	21.93 (13.67, 39.02)
		(M) 3 best reps.	0.86 (0.46, 0.95)	6.44 (3.85, 12.65)	17.85 (10.67, 35.06)
IRinHF@45°	Affected	Best of 3 reps.	0.93 (0.84, 0.97)	15.51 (10.15, 23.45)	42.99 (28.15, 65.00)
		(M) 3 best reps.	0.94 (0.84, 0.97)	12.74 (9.01, 20.81)	35.32 (24.98, 57.69)
	Unaffected	Best of 3 reps.	0.92 (0.83, 0.97)	18.20 (11.15, 26.54)	50.45 (30.90, 73.55)
		(M) 3 best reps.	0.94 (0.85, 0.97)	14.37 (10.16, 22.73)	39.84 (28.17, 62.99)

ER@0°: external rotation at 0° of shoulder abduction; IR@0°: internal rotation at 0° of shoulder abduction; ER@90°: external rotation at 90° of shoulder abduction; IR@90°: internal rotation at 90° of shoulder abduction; Abd@45°: abduction in 45° of shoulder abduction; ERinHF@45°: external rotation in 45° of horizontal flexion; IRinHR@45°: internal rotation in 45° of horizontal flexion, Ext@90°: extension in 90° of shoulder abduction; Flex@90°: flexion at 90° of shoulder abduction.
ER: external rotation; IR: internal rotation; HF: horizontal flexion; ICC: intra-class correlation coefficient; SEM: standard error of measurement; MDC: minimal detectable change; M: mean; Reps: repetitions

both raters. This suggests that the HHD protocol described in this study, when used by an experienced clinician, has clinical applicability for both asymptomatic populations and shoulders with symptomatic instability.

Despite standardised testing procedures, a rater's level of experience may influence reliability of HHD as a measure of strength. In our study, each therapist had over 10 years of clinical experience, working exclusively in shoulder assessment and rehabilitation and have previous experience using HHD. Consequently, there may be limitations in the translation of study findings to less experienced clinicians. Future research is needed to explore the difference between raters with differing clinical expertise, sex and anthropometrics.

Limitations

Participant strength measures were collected by using pragmatic methods to reflect the clinical environment. For scheduling purposes, participants had one submaximal practice test prior to the collection of data, resulting in possible learning effects, or self-determined pacing. An operational decision was made to un-blind therapists to the side of shoulder instability, to mitigate potential adverse events (such as subluxation), though therapists used the same strength assessment protocol in both asymptomatic and symptomatic shoulders. It is also important to note that limb dominance was not considered during testing sessions. It is unclear what effect this had on the strength assessments completed and the study findings.

When measuring strength using a HHD, a 'make test' or a 'break test' have both been shown to be reliable.[47] Higher force values may be obtained with a 'break test' protocol[43]; however, this form of testing requires the therapist to resist against the participant's limb until the subject's maximal muscular effort is overcome and the joint being tested gives way.[47] The decision to use a 'make test' protocol, where the therapist remains still during the participants force exertion,[47] was made a priori to mitigate any potential adverse risks that may be associated with a 'break test' protocol in a clinically unstable shoulder population. Though, using the aforementioned methods, this study identified a ceiling effect in reliability for positions where force output was >300N, such as the shrug test.

Conclusion

This study has demonstrated that shoulder strength measured using a portable HHD by experienced therapists with pragmatic methods has moderate to high reliability in student circus performers. This study supports use of the HHD in the clinical environment to test shoulder strength in both asymptomatic (control) and symptomatic shoulders with atraumatic shoulder instability. To optimise reliability, testing should be completed using consistent positions with the same therapist. The use of HHD may facilitate enhanced clinical strength assessment procedures, guide interventions, and return to sport in people with atraumatic shoulder instability.

Acknowledgments: The authors thank The National Institute of Circus Arts (NICA) for approving the participation of their student circus performers in this study.

Authors' Contributions: CG, SW, RL, DM, SM, LW and OT designed and implemented the study and collected data. MK, OT and BK completed data analysis. All authors constructed, edited, and approved the final version of the manuscript.

REFERENCES

1. Munro D. Injury patterns and rates amongst students at the national Institute of Circus Arts: an observational study. *Med Probl Perform Art*. 2014;29(4):235–40. https://doi.org/10.21091/mppa.2014.4046
2. Stubbe JH, Richardson A, van Rijn RM. Prospective cohort study on injuries and health problems among circus arts students. *BMJ Open Sport Exerc Med*. 2018;4(1). https://doi.org/10.1136/bmjsem-2017-000327
3. Shrier I, Meeuwisse WH, Matheson GO, et al. Injury patterns and injury rates in the circus arts: an analysis of 5 years of data from Cirque du Soleil. *Am J Sports Med*. 2009;37(6):1143–9. https://doi.org/10.1177/0363546508331138
4. Cadogan A, Laslett M, Hing W, et al. Reliability of a new hand-held dynamometer in measuring shoulder range of motion and strength. *Man Ther*. 2011;16:97–101. https://doi.org/10.1016/j.math.2010.05.005
5. Beshay N, Lam PH, Murrell GAC. Assessing the reliability of shoulder strength measurement: hand-held versus fixed dynamometry. *Shoulder Elbow*. 2011;3(4):244–51. https://doi.org/10.1111/j.1758-5740.2011.00137.x
6. Cools AM, De Wilde L, Van Tongel A, Cet al. Measuring shoulder external and internal rotation strength and range of motion: Comprehensive intra-rater and inter-rater reliability study of several testing protocols. *J Shoulder Elbow Surg*. 2014;23(10):1454–61. https://doi.org/10.1016/j.jse.2014.01.006
7. Conceição A, Parraca J, Marinho D, et al. Assessment of isometric strength of the shoulder rotators in swimmers using a hand-held dynamometer: a reliability study. *Acta Bioeng Biomech*. 2018;20(4):113–9. https://doi.org/10.5277/ABB-01237-2018-02
8. Dollings H, Sandford F, O'Conaire E, Lewis JS. Shoulder strength testing: the intra- and inter-tester reliability of routine clinical tests, using the PowertrackTM II Commander. *Shoulder Elbow*. 2012;4(2):131–40. https://doi.org/10.1111/j.1758-5740.2011.00162.x
9. Karabay D, Yesilyaprak SS, Sahiner Picak G. Reliability and validity of eccentric strength measurement of the shoulder abductor muscles using a hand-held dynamometer. *Phys Ther Sport*. 2020;43:52–7. https://doi.org/10.1016/j.ptsp.2020.02.002
10. Katoh M. Test-retest reliability of isometric shoulder muscle strength measurement with a handheld dynamometer and belt. *J Phys Ther Sci*. 2015;27(6):1719–22. https://doi.org/10.1589/jpts.27.1719
11. Vermeulen HM, de Bock GH, van Houwelingen HC, et al. A comparison of two portable dynamometers in the assessment of shoulder and elbow strength. *Physiother*. 2005;91(2):101–12. https://doi.org/10.1016/j.physio.2004.08.005
12. Holt KL, Raper DP, Boettcher CE, et al. Hand-held dynamometry strength measures for internal and external rotation demonstrate superior reliability, lower minimal detectable change and higher correlation to isokinetic dynamometry than externally-fixed dynamometry of the shoulder. *Phys Ther Sport*. 2016;21(75–81):75–81. https://doi.org/10.1016/j.ptsp.2016.07.001
13. Awatani T, Mori S, Shinohara J, et al. Same-session and between-day intra-rater reliability of hand-held dynamometer measurements of isometric shoulder extensor strength. *J Phys Ther Sci*. 2016;28(3):936–9. https://doi.org/10.1589/jpts.28.936
14. McCormack M, Briggs J, Hakim A, Grahame R. Joint laxity and the benign joint hypermobility syndrome in student and professional ballet dancers. *J Dance Med Sci*. 2004;8:89.
15. Hinds N, Angioi M, Birn-Jeffery A, Twycross-Lewis R. A systematic review of shoulder injury prevalence, proportion, rate, type, onset, severity, mechanism and risk factors in female artistic gymnasts. *Phys Ther Sport*. 2019;35:106–15. https://doi.org/10.1016/j.ptsp.2018.11.012
16. Burnham SR, Bell CG, Olenik CL, Reid CD. Shoulder abduction strength measurement in football players: Reliability and validity of two field tests. *Clin J Sport Med*. 1995;5(2):90–4. https://doi.org/10.1097/00042752-199504000-00004
17. Martin HJ, Yule V, Syddall HE, et al. Is hand-held dynamometry useful for the measurement of quadriceps strength in older people? A comparison with the gold standard Bodex dynamometry. *Gerontology*. 2006;52(3):154–9. https://doi.org/10.1159/000091824
18. Lund I, Lundeberg T, Sandberg L, et al. Lack of interchangeability between visual analogue and verbal rating pain scales: a cross sectional description of pain etiology groups. *BMC Med Res Methodol*. 2005;5:31. https://doi.org/10.1186/1471-2288-5-31
19. Stark T, Walker B, Phillips JK, et al. Hand-held dynamometry correlation with the gold standard isokinetic dynamometry: a systematic review. *PM&R*. 2011;3(5):472–9. https://doi.org/10.1016/j.pmrj.2010.10.025
20. Bak K. Nontraumatic glenohumeral instability and coracoacromial impingement in swimmers. *Scand J Med Sci Sport*. 1996;6(3):132–44. https://doi.org/10.1111/j.1600-0838.1996.tb00081.x
21. Zemek M, Magee D. Comparison of glenohumeral joint laxity in elite and recreational swimmers. *Clin J Sport Med*. 1996;6:40–7. https://doi.org/10.1097/00042752-199601000-00009
22. Schmidt H, Pedersen TL, Junge T, et al. Hypermobility in adolescent athletes: pain, functional ability, quality of life, and musculoskeletal injuries. *J Orthop Sports Phys Ther*. 2017;47(10):792–800.
23. Warby SA, Watson L, Ford JJ, et al. Multidirectional instability of the glenohumeral joint: Etiology, classification, assessment, and management. *J Hand Ther*. 2017;30(2):175–81. https://doi.org/10.1016/j.jht.2017.03.005
24. Thomas SC, Matsen FA. An approach to the repair of avulsion of the glenohumeral ligaments in the management of traumatic anterior glenohumeral instability. *J Bone Joint Surg Am*. 1989;71(4):506–13. Epub 1989/04/01. PubMed PMID: 2703510.
25. Bahk M, Keyurapan E, Tasaki A, et al. Laxity testing of the shoulder: a review. *Am J Sports Med*. 2007;35(1):131–44.

https://doi.org/10.1177/0363546506294570

26. Tzannes A, Paxinos A, Callanan M, Murrell GAC. An assessment of the interexaminer reliability of tests for shoulder instability. *J Shoulder Elbow Surg.* 2004;13(1):18–23. https://doi.org/10.1016/j.jse.2003.09.002
27. McFarland EG. *Instability and Laxity*. New York: Thieme; 2006.
28. Levy AS, Lintner S, Kenter K, Speer KP. Intra- and interobserver reproducibility of the shoulder laxity examination. *Am J Sports Med*. 1999;27(4):460.
29. Gerber C, Ganz R. Clinical assessment of instability of the shoulder. With special reference to anterior and posterior drawer tests. *J Bone Joint Surg Br*. 1984;66(4):551–6. https://doi.org/10.1302/0301-620X.66B4.6746691
30. Rowe CR, Zarins B. Recurrent transient subluxation of the shoulder. *J Bone Joint Surg Am*. 1981;63:863–72.
31. Rowe C, Leffert RD. *The Shoulder*. New York: Churchhill Livingstone; 1988: p 165–291.
32. Kessell L. *Clinical Disorders of the Shoulder*. Churchill Livingstone; 1981: p 196.
33. Eshoj H, Ingwersen KG, Larsen CM, et al. Intertester reliability of clinical shoulder instability and laxity tests in subjects with and without self-reported shoulder problems. *BMJ Open*. 2018; 8(3):e018472. https://doi.org/https://doi.org/10.1136/bmjopen-2017-018472
34. Beighton P, Horan F. Orthopaedic aspects of the Ehlers-Danlos syndrome. *J Bone Joint Surg Br*. 1969;51(3):444–53. Epub 1969/08/01. PubMed PMID: 5820785.
35. Anekstein Y, Blecher R, Smorgick Y, Mirovsky Y. What is the best way to apply the Spurling test for cervical radiculopathy? *Clin Orthop Relat Res*. 2012(9):2566–72. https://doi.org/doi:10.1007/s11999
36. Warby SA, Ford JJ, Hahne AJ, et al. Comparison of 2 exercise rehabilitation programs for multidirectional instability of the glenohumeral joint: a randomized controlled trial. *Am J Sports Med*. 2018;46(1):87–97. https://doi.org/10.1177/0363546517734508
37. Thorborg K, Bandholm T, Schick M, et al. Hip strength assessment using handheld dynamometry is subject to intertester bias when testers are of different sex and strength. *Scand J Med Sci Sports*. 2013;23(4):487–93. https://doi.org/10.1111/j.1600-0838.2011.01405.x
38. Koo TK, Li MY. A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *J Chiropract Med*. 2016;15(2):155–63. https://doi.org/10.1016/j.jcm.2016.02.012
39. Weir PJ. Quantifying test-retest reliability using the intraclass correlation coefficient and the SEM. *J Strength Cond Res*. 2005; 19(1):231–40. https://doi.org/10.1519/15184.1
40. Meyers CR. *Measurement in Physical Education*. New York: Ronald Press Co.; 1974.
41. Meeteren J, Roebroeck M, Stam H. Test-retest reliability in isokinetic muscle strength measurements of the shoulder. *J Rehabil Med*. 2002;34:91–5. https://doi.org/10.1080/165019702753557890
42. Hirano M, Katoh M. Limits of the manipulative-fixed method for measurement of shoulder joint horizontal adduction muscle strength using a handheld dynamometer. *J Phys Ther Sci*. 2015; 27(1):235–7. https://doi.org/10.1589/jpts.27.235
43. Thorborg K, Petersen J, Magnusson SP, Hölmich P. Clinical assessment of hip strength using a hand held dynamometer is reliable. *Scand J Med Sci Sports*. 2010;20(3):493–501. https://doi.org/10.1111/j.1600-0838.2009.00958.x
44. Wolf MJ, Cameron LK, Owens DB. Impact of joint laxity and hypermobility on the musculoskeletal system. *J Am Acad Orthop Surg*. 2011;19(8):463–71. https://doi.org/10.5435/00124635-201108000-00002
45. Wanke EM, McCormack M, Koch F, et al. Acute injuries in student circus artists with regard to gender specific differences. *Asian J Sports Med*. 2012;3(3):153. https://doi.org/10.5812/asjsm.34606
46. Pizzari T, Wickham J, Watson L, et al. Altered muscle activation patterns in multidirectional shoulder instability during dynamic abduction. *J Sci Med Sport*. 2009;12:S4–S. https://doi.org/10.1016/j.jsams.2008.12.011
47. Bohannon RW. Make tests and break tests of elbow flexor muscle strength. *Phys Ther Sport*. 1988;68(2):193–4.

Received 22-Dec-2020, accepted 22-Apr-2021
Published online 1-June-2021
https://doi.org/10.21091/mppa.2021.2014

APPENDIX A. Glenohumeral Joint Instability Assessment

1. **Sulcus Sign**[21,23,27]

   Description: Performed with patient standing or sitting relaxed with the shoulder in the neutral position, the therapist stabilises the scapula and places an inferior distraction force to the upper arm in an inferior direction.

2. **Anterior and posterior draw test**
   (10–30° abduction)[21,23,27]

   Description: Performed with patient supine, the examiner stands to the affected side and supports the hand in their axilla by adducting their humerus. The patient's arm is placed in 10–30° abduction, 20–30° forward flexion and 0–30° of external rotation. They place one hand on the spine of the scapular and coracoid process and the other hand on the humeral shaft to provide an anteriorly or posteriorly directed force to assess the amount of humeral head translation.

3. **Anterior and posterior drawer test abducted**
   (80–120° abduction)[21,23,27]

   Description: Performed with patient supine, the examiner stands to the affected side and supports the hand in their axilla by adducting their humerus. The patient's arm is placed in 80–120° abduction, 20–30° forward flexion and 0–30° of external rotation. They place one hand on the spine of the scapular and coracoid process and the other hand on the humeral shaft to provide an anteriorly or posteriorly directed force to assess the amount of humeral head translation.

4. **Anterior and posterior apprehension test**[21,24]

   Anterior Apprehension Description: Therapist externally rotates the humerus in 45°, 90° and 135° of abduction with one hand, combined with forward pressure on the humeral head with the other hand to identify reported (subjective/ involuntary physical) apprehension.

   Posterior Apprehension Description: Performed with patient in supine, the therapist guides the arm into 90° horizontal flexion and internal rotation. Whilst palpating the posterior humeral head, the therapist applies a posterior force to the lower arm to identify reported (subjective / involuntary physical) apprehension.

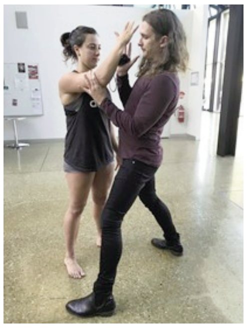
5. **Active movement dysfunction**[21]

   Description: Using visual observation and palpation of the humeral head during the following active movements: abduction, flexion, horizontal flexion with or without internal rotation, horizontal extension with external rotation. The therapist aims to identify displacement or loss of centering of the humeral head within the glenoid.

APPENDIX B. Hand-Held Dynamometer Testing Positions

Transducer head aligned just proximal to the ulnar styloid process.

	1. **External rotation:** The subject was standing with feet separated at shoulder width, arm kept in a neutral position by the side but not touching the body, with their elbow bent at 90° flexion and wrist in anatomical neutral with the palm facing midline. The dynamometer was held parallel to the floor on the volar aspect of the distal forearm. The rater stood to the affected side with one hand on the dynamometer and one hand on the anterior and posterior joint line of the shoulder to assess for humeral head translation.
	2. **Internal rotation:** The subject was standing with feet separated at shoulder width, arm kept in a neutral position by the side but not touching the body, with their elbow bent at 90° flexion and wrist in anatomical neutral with the palm facing midline. The dynamometer was held parallel to the floor on the palmar aspect of the distal forearm. The rater stood to the unaffected side with one hand on the dynamometer and one hand on the anterior and posterior joint line of the shoulder to assess for humeral head translation.
	3. **External rotation at 90°:** The subject was standing with feet separated at shoulder width in a wide stance with slightly flexed knees and hips. Arm was held at 90° abduction, external rotation to 90° and 90° elbow flexion with the wrist in anatomical neutral with the palm facing midline. The dynamometer was held parallel to the floor on the volar aspect of the distal forearm. The rater stood behind the subject, to the affected side with one hand on the dynamometer and one hand on the anterior and posterior joint line of the shoulder to assess for humeral head transition.
	4. **Internal rotation at 90°:** The subject was standing with feet separated at shoulder width in a wide stance with slightly flexed knees and hips. Arm was held at 90° abduction, external rotation to 90° and 90° elbow flexion with the wrist in anatomical neutral with the palm facing forwards. The dynamometer was held parallel to the floor on the palmar aspect of the distal forearm. The rater stood in front of the subject to the affected side with one hand on the dynamometer and one hand on the anterior and posterior joint line of the shoulder to assess for humeral head translation.
	5. **Abduction at 45°:** The subject was standing with feet separated at shoulder width. Arm was held at 45° abduction, neutral rotation and 90° elbow flexion with the wrist in anatomical neutral with the palm facing forwards. The dynamometer was held 45° to the floor on the lateral aspect of the distal arm. The rater stood to the affected side with one hand on the dynamometer and one hand on the anterior and posterior joint line of the shoulder to assess for humeral head translation.

	6. **External rotation in horizontal flexion at 90°:** The subject was standing with feet separated at shoulder width. Arm was held at 90° flexion, internal rotation to 45° and 90° elbow flexion with the wrist in anatomical neutral with the midline. The dynamometer was held 45° to the floor on the volar aspect of the distal forearm. The rater stood to the affected side with one hand on the dynamometer and one hand stabilising the elbow.
	7. **Internal rotation in horizontal flexion at 90°:** The subject was standing with feet separated at shoulder width. Arm was held at 90° flexion, internal rotation to 45° and 90° elbow flexion with the wrist in anatomical neutral with the midline. The dynamometer was held 45° to the floor on the palmar aspect of the distal forearm. The rater stood to the unaffected side with one hand on the dynamometer and one hand stabilising the elbow.
	8. **Extension at 90°:** The subject was lying prone. Arm was held at 90° abduction, neutral rotation, 90° elbow flexion with the wrist in anatomical neutral. The dynamometer was held perpendicular to the floor on the posterior aspect of the distal arm. The rater stood to the affected side with one hand on the dynamometer and one hand stabilising the torso.
	9. **Flexion at 90°:** The subject was kneeling with one the opposite foot forward. Arm was held at 90° flexion, neutral rotation, 90° elbow flexion with the wrist in anatomical neutral and hand in a flexed fist. The dynamometer was held perpendicular to the floor on the fist. The rater stood to the affected side with both hands on the dynamometer.
	10. **Shrug:** The subject was standing with feet separated at shoulder width. Arm was held at 30° abduction, arm and wrist in anatomical neutral with the palm facing forwards. The dynamometer was fixed via a belt from the top of the scapular and under the foot. The rater stood to the affected side stabilising the dynamometer with one hand.

APPENDIX C. Intra-rater Reliability

	Rater	Arm		Test (N)*	Retest (N)*	Difference	Paired t-Test	$ICC_{3,1}$ (95%CI) (intra-rater reliability)	SEM (95%CI) Test-Re-Test	%SEM	MDC (95%CI)
ER@0°	1	Affected	Best of 3 reps.	90.91 (22.82)	94.77 (24.86)	−3.86 (14.45)	0.214	0.90 (0.76, 0.96)	4.57 (2.89, 7.08)	5%	12.67 (8.01, 19.62)
			(M) 3 best reps.	85.78 (22.18)	87.61 (21.17)	−1.83 (11.77)	0.463	0.92 (0.81, 0.97)	3.33 (2.04, 5.13)	4%	9.23 (5.65, 14.22)
		Unaffected	Best of 3 reps.	92.44 (25.83)	90.32 (22.41)	2.13 (11.77)	0.396	0.94 (0.85, 0.97)	2.88 (2.04, 4.56)	3%	7.99 (5.65, 12.64)
			(M) 3 best reps.	86.90 (24.00)	84.26 (20.64)	2.65 (11.57)	0.285	0.93 (0.83, 0.97)	3.06 (2.00, 4.77)	4%	8.49 (5.55, 13.22)
	2	Affected	Best of 3 reps.	97.86 (20.24)	106.95 (25.49)	−9.09 (11.87)	0.001	0.92 (0.83, 0.97)	3.36 (2.06, 4.89)	3%	9.31 (5.70, 13.57)
			(M) 3 best reps.	93.36 (19.38)	100.76 (23.46)	−7.41 (10.57)	0.003	0.94 (0.85, 0.97)	2.59 (1.83, 4.09)	3%	7.18 (5.07, 11.35)
		Unaffected	Best of 3 reps.	99.99 (21.82)	107.15 (23.18)	−7.16 (8.03)	<0.001	0.97 (0.92, 0.99)	1.39 (0.8, 2.27)	1%	3.86 (2.23, 6.30)
			(M) 3 best reps.	95.67 (20.97)	102.18 (22.27)	−6.51 (7.30)	<0.001	0.97 (0.93, 0.99)	1.26 (0.73, 1.93)	1%	3.50 (2.02, 5.35)
IR@0°	1	Affected	Best of 3 reps.	148.71 (49.78)	136.92 (45.51)	11.79 (29.78)	0.071	0.89 (0.75, 0.95)	9.88 (6.66, 14.89)	7%	27.38 (18.46, 41.27)
			(M) 3 best reps.	137.70 (47.96)	126.22 (42.72)	11.48 (24.87)	0.038	0.92 (0.81, 0.97)	7.03 (4.31, 10.84)	5%	19.50 (11.94, 30.05)
		Unaffected	Best of 3 reps.	143.99 (48.35)	133.44 (40.35)	10.55 (24.91)	0.055	0.92 (0.80, 0.96)	7.05 (4.98, 11.14)	5%	19.53 (13.81, 30.88)
			(M) 3 best reps.	134.65 (46.22)	125.06 (38.85)	9.59 (21.31)	0.042	0.93 (0.84, 0.97)	5.64 (3.69, 8.52)	4%	15.63 (10.23, 23.63)
	2	Affected	Best of 3 reps.	139.43 (42.47)	149.30 (39.57)	−9.87 (16.20)	0.008	0.96 (0.90, 0.98)	3.24 (2.29, 5.12)	2%	8.98 (6.35, 14.20)
			(M) 3 best reps.	132.15 (40.07)	142.14 (37.76)	−9.99 (18.36)	0.016	0.94 (0.86, 0.98)	4.50 (2.60, 6.87)	3%	12.47 (7.20, 19.04)
		Unaffected	Best of 3 reps.	139.43 (45.66)	151.03 (39.85)	−11.60 (16.07)	0.003	0.96 (0.91, 0.99)	3.21 (1.61to 4.82)	2%	8.91 (4.45, 13.36)
			(M) 3 best reps.	130.67 (40.36)	142.91 (38.63)	−12.24 (14.77)	0.001	0.96 (0.92, 0.99)	2.95 (2.95, 4.18)	2%	8.19 (8.19, 11.58)
ER@90°	1	Affected	Best of 3 reps.	69.07 (19.80)	61.70 (14.78)	7.37 (10.26)	0.002	0.91 (0.78, 0.96)	3.08 (2.05, 4.81)	4%	8.53 (5.69, 13.34)
			(M) 3 best reps.	63.82 (18.60)	55.62 (13.76)	8.20 (10.26)	0.001	0.89 (0.74, 0.95)	3.40 (2.29, 5.23)	5%	9.43 (6.36, 14.50)
		Unaffected	Best of 3 reps.	69.43 (23.12)	61.70 (16.62)	7.73 (15.49)	0.026	0.82 (0.59, 0.93)	6.57 (4.10, 9.92)	10%	18.22 (11.36, 27.49)
			(M) 3 best reps.	63.50 (19.62)	57.05 (15.06)	6.44 (11.71)	0.015	0.87 (0.70, 0.95)	4.22 (2.62, 6.41)	7%	11.70 (7.26, 17.79)
	2	Affected	Best of 3 reps.	64.20 (14.19)	82.96 (21.65)	−18.76 (11.38)	<0.001	0.89 (0.75, 0.96)	3.77 (2.28, 5.69)	6%	10.46 (6.31, 15.77)
			(M) 3 best reps.	60.27 (13.47)	77.99 (20.05)	−17.72 (9.71)	<0.001	0.91 (0.79, 0.96)	2.91 (1.94, 4.45)	5%	8.07 (5.38, 12.33)
		Unaffected	Best of 3 reps.	65.96 (13.33)	84.72 (18.61)	−18.76 (10.64)	<0.001	0.88 (0.71, 0.95)	3.69 (2.38, 5.73)	6%	10.22 (6.59, 15.88)
			(M) 3 best reps.	61.07 (13.45)	79.43 (18.11)	−18.36 (10.06)	<0.001	0.89 (0.74, 0.95)	3.34 (2.25, 5.13)	5%	9.25 (6.24, 14.22)
IR@90°	1	Affected	Best of 3 reps.	96.12 (41.23)	92.56 (33.57)	3.56 (28.21)	0.551	0.84 (0.61, 0.93)	11.28 (7.46, 17.62)	12%	31.28 (20.69, 48.83)
			(M) 3 best reps.	86.88 (34.73)	85.03 (31.08)	1.84 (24.35)	0.720	0.84 (0.62, 0.93)	9.74 (6.44, 15.01)	11%	27.00 (17.86, 41.61)
		Unaffected	Best of 3 reps.	108.11 (47.72)	100.75 (40.57)	7.36 (32.32)	0.287	0.85 (0.64, 0.94)	12.52 (7.92, 19.39)	12%	34.70 (21.94, 53.75)
			(M) 3 best reps.	98.97 (42.35)	90.38 (32.89)	8.60 (29.26)	0.173	0.83 (0.59, 0.93)	12.06 (7.74, 18.74)	12%	33.44 (21.46, 51.93)
	2	Affected	Best of 3 reps.	97.67 (32.06)	122.42 (39.41)	−24.75 (17.52)	<0.001	0.94 (0.85, 0.97)	4.29 (3.03, 4.29)	4%	11.90 (8.41, 11.90)
			(M) 3 best reps.	92.31 (31.48)	114.64 (36.34)	−22.33 (16.98)	<0.001	0.93 (0.84, 0.97)	4.49 (2.94, 4.49)	5%	12.45 (8.15, 12.45)
		Unaffected	Best of 3 reps.	96.32 (32.53)	121.65 (34.08)	−25.33 (16.87)	<0.001	0.93 (0.84, 0.97)	4.46 (2.92, 6.75)	5%	12.37 (8.10, 18.70)
			(M) 3 best reps.	91.22 (31.80)	114.10 (31.80)	−22.88 (15.35)	<0.001	0.94 (0.85, 0.97)	3.76 (3.76, 5.95)	4%	10.42 (10.42, 16.48)

(continued on next page)

APPENDIX C. Intra-rater Reliability (*continued*)

	Rater	Arm		Test (N)*	Retest (N)*	Difference	Paired *t*-Test	$ICC_{3,1}$ (95%CI) (intra-rater reliability)	SEM (95%CI) Test-Re-Test	%SEM	MDC (95%CI)
Shrug	1	Affected	Best of 3 reps.	338.97 (101.95)	406.77 (138.98)	−67.91 (88.87)	0.001	0.85 (0.64, 0.94)	34.42 (21.77, 53.32)	10%	95.41 (60.34, 147.80)
			(M) 3 best reps.	305.47 (92.65)	370.97 (131.24)	−65.51 (86.70)	0.002	0.83 (0.60, 0.93)	35.75 (22.94, 54.83)	12%	99.09 (63.58, 151.99)
		Unaffected	Best of 3 reps.	348.64 (111.85)	401.67 (144.00)	−53.06 (100.35)	0.019	0.82 (0.58, 0.92)	46.82 (31.21, 71.51) 13%/11%	13%	129.77 (86.51, 198.23)
			(M) 3 best reps.	311.02 (101.03)	368.25 (137.95)	−57.23 (93.31)	0.008	0.83 (0.59, 0.93)	38.47 (24.69, 59.75)	12%	106.64 (68.43, 165.61)
	2	Affected	Best of 3 reps.	384.45 (117.98)	424.47 (139.59)	−40.02 (59.65)	0.004	0.94 (0.87, 0.98)	14.61 (8.44, 21.51)	4%	40.50 (23.38, 59.61)
			(M) 3 best reps.	356.42 (110.68)	390.39 (127.45)	−33.97 (55.22)	0.007	0.94 (0.87, 0.98)	13.53 (7.81, 19.91)	4%	37.49 (21.65, 55.19)
		Unaffected	Best of 3 reps.	403.13 (141.48)	421.40 (137.75)	−18.27 (53.54)	0.116	0.96 (0.91, 0.98)	10.71 (7.57, 16.06)	3%	29.68 (20.99, 44.52)
			(M) 3 best reps.	364.00 (126.55)	387.98 (126.03)	−23.99 (62.18)	0.073	0.94 (0.85, 0.97)	15.23 (10.77, 24.08)	4%	42.22 (29.85, 66.75)
ERinHF@45°	1	Affected	Best of 3 reps.	61.12 (22.06)	56.48 (20.97)	4.64 (15.02)	0.153	0.86 (0.67, 0.94)	5.62 (3.68, 8.63)	9%	15.58 (10.20, 23.92)
			(M) 3 best reps.	55.26 (19.09)	52.83 (19.90)	2.43 (11.15)	0.307	0.91 (0.79, 0.96)	3.35 (2.23, 5.11)	6%	9.27 (6.18, 14.16)
		Unaffected	Best of 3 reps.	57.52 (16.37)	59.96 (17.27)	−2.67 (10.13)	0.219	0.90 (0.77, 0.96)	3.20 (2.03, 4.86)	6%	8.88 (5.62, 13.47)
			(M) 3 best reps.	52.64 (13.54)	53.57 (15.86)	−0.93 (9.61)	0.648	0.88 (0.72, 0.95)	3.33 (2.15, 5.09)	6%	9.23 (5.96, 14.10)
	2	Affected	Best of 3 reps.	62.67 (18.74)	76.79 (20.17)	−14.12 (8.97)	<0.001	0.94 (0.87, 0.98)	2.20 (1.27, 3.23)	4%	6.09 (3.52, 8.96)
			(M) 3 best reps.	59.37 (17.51)	72.14 (19.10)	−12.77 (8.67)	<0.001	0.94 (0.86, 0.98)	2.12 (1.23, 3.24)	4%	5.89 (3.40, 8.99)
		Unaffected	Best of 3 reps.	65.76 (17.74)	78.71 (18.80)	−12.96 (12.29)	<0.001	0.87 (0.70, 0.95)	4.43 (2.75, 6.73)	7%	12.28 (7.62, 18.66)
			(M) 3 best reps.	60.02 (16.36)	73.88 (17.75)	−13.87 (10.63)	<0.001	0.89 (0.75, 0.95)	3.53 (2.38, 5.32)	6%	9.77 (6.59, 14.73)
IRinHF@45°	1	Affected	Best of 3 reps.	140.65 (55.45)	147.75 (64.63)	−7.10 (35.43)	0.347	0.91 (0.78, 0.96)	10.63 (7.09 to16.62)	8%	29.46 (19.64, 46.06)
			(M) 3 best reps.	130.20 (49.02)	134.92 (58.72)	−4.72 (32.75)	0.497	0.91 (0.76, 0.96)	9.83 (6.55, 16.04)	8%	27.23 (18.16, 44.47)
		Unaffected	Best of 3 reps.	140.40 (73.00)	145.62 (61.65)	−5.22 (35.98)	0.494	0.92 (0.82, 0.97)	10.11 (6.19, 15.17)	7%	28.03 (17.16, 42.04)
			(M) 3 best reps.	129.99 (65.95)	130.36 (50.71)	−0.37 (31.13)	0.955	0.96 (0.90, 0.98)	6.23 (9.84, 4.40)	5%	17.26 (12.20, 27.29)
	2	Affected	Best of 3 reps.	151.23 (50.73)	166.31 (48.63)	−15.08 (21.44)	0.003	0.96 (0.91, 0.98)	4.29 (3.03, 6.43)	3%	11.89 (8.40, 17.83)
			(M) 3 best reps.	139.68 (52.36)	158.57 (46.07)	−18.90 (19.75)	<0.001	0.92 (0.82, 0.96)	5.59 (3.95, 8.38)	4%	15.48 (10.95, 23.23)
		Unaffected	Best of 3 reps.	147.74 (54.56)	166.89 (49.15)	−19.14 (22.17)	<0.001	0.95 (0.89, 0.98)	4.96 (3.14, 7.35)	3%	13.74 (8.69, 20.38)
			(M) 3 best reps.	139.31 (59.91)	157.29 (45.91)	−17.98 (19.83)	<0.001	0.96 (0.90, 0.98)	3.97 (2.80, 6.27)	3%	10.99 (7.77, 17.38)

* For Test (N), Retest (N), and Difference day 1–4, data given as mean (SD). For SEM and MDC, data given in Newtons. ER, external rotation; IR, internal rotation; HF, horizontal flexion; ICC, intra-class correlation coefficient; SEM, standard error of measurement; MDC, minimal detectable change; M, mean; reps., repetitions.

ER@0°: external rotation at 0° of shoulder abduction; IR@0°: internal rotation at 0° of shoulder abduction; ER@90°: external rotation at 90° of shoulder abduction; IR@90°: internal rotation at 90° of shoulder abduction; Abd@45°: abduction in 45° of shoulder abduction; ERinHF@45°: external rotation in 45° of horizontal flexion; IRinHR@45°: internal rotation in 45° of horizontal flexion, Ext@90°: extension in 90° of shoulder abduction; Flex@90°: flexion at 90° of shoulder abduction.