Central European Journal of Sport Sciences and Medicine | Vol. 38, No. 2/2022: 47–59
ISSN (print): 2300-9705 | ISSN (online): 2353-2807 | DOI: 10.18276/cej.2022.2-05

RELIABILITY OF THE COACH'S EYE GONIOMETER APPLICATION DURING SQUAT EXERCISE

Jonathan Kilgallon[A, B, C, D]

Faculty of Sport, Allied Health and Performance Science, St Mary's University, Twickenham, United Kingdom
ORCID: 0000-0003-2404-1950 | jkperformance.info@gmail.com

Joseph Cave[B]

Faculty of Sport, Allied Health and Performance Science, St Mary's University, Twickenham, United Kingdom

Emily Cushion[D]

Faculty of Sport, Allied Health and Performance Science, St Mary's University, Twickenham, United Kingdom
ORCID: 0000-0001-7582-9575

Shaun Joffe[D]

Faculty of Sport, Allied Health and Performance Science, St Mary's University, Twickenham, United Kingdom
British Weightlifting, Leeds, United Kingdom
ORCID: 0000-0002-2234-6278

Jamie Tallent[A, C ,D]

School of Sport, Rehabilitation and Exercise Sciences, University of Essex, Colchester, United Kingdom
Department of Physiotherapy, School of Primary and Allied Health Care, Faculty of Medicine Nursing and Health Science, Monash University, Melbourne, Australia
ORCID: 0000-0002-4354-9912

[A] Study Design; [B] Data Collection; [C] Statistical Analysis; [D] Manuscript preparation; [E] Funds Collection.

Abstract This study examined the test re-test, intrarater and interrater reliability of joint kinematics from the Coach's Eye smartphone application. Twenty-two males completed a 1-repetition maximum (1-RM) assessment followed by 2 identical sessions using 5 incremental loads (20, 40, 60, 80, 90% 1-RM). Peak flexion angles at the hip, knee, and ankle joints were assessed using 1 experienced practitioner and 1 inexperienced practitioner. The acceptable reliability thresholds were defined as intraclass correlation coefficient (ICC) $r > 0.70$ and coefficient of variation $CV \leq 10\%$. The test re-test reliability of peak hip and knee flexion were reliable across 20–90% 1-RM ($r > 0.64$; $CV < 4.2\%$), whereas peak ankle flexion was not reliable at any loaded condition ($r > 0.70$; $CV < 20.4\%$). No significant differences were detected between trials ($p > 0.11$). The intrarater reliability was near perfect ($r > 0.90$) except for peak ankle flexion ($r > 0.85$). The interrater reliability was nearly perfect ($r > 0.91$) except for hip flexion at 80% 1-RM and ankle flexion at 20% ($r > 0.77$). Concludingly, the Coach's Eye application can produce repeatable assessments of joint kinematics using either a single examiner or 2 examiners, regardless of experience level. The Coach's Eye can accurately monitor squat depth.

Key words range of motion, kinematics, lower limb joints, two-dimensional analysis, rehabilitation

Introduction

The back squat is a closed kinetic chain exercise requiring coordination at the hip, knee, and ankle (Schoenfeld, 2010). The back squat is commonly used by practitioners in rehabilitation and strength and conditioning (S&C) programs to assess an individual's neuromuscular control, strength, stability, and mobility within the kinetic chain (Escamilla et al., 1998; Hartmann et al., 2012; Myer, Paterno, Ford, Hewett, 2008; Wirth et al., 2016). The reliable and valid assessment of back squat mechanics provides useful information for S&C coaches and physical therapists regarding an individual's functional capacities or risk of injury. For instance, variation in squat depth is known to influence the development of kinetic and kinematic outcomes (Martinez-Cava, Moran-Navarro, Sanchez-Medina, Gonzalez-Badillo, Pallares, 2019; Rhea et al., 2016). While abnormal lower extremity kinematics during a deep squat may infer movement limitations stemming from mobility issues (Kim, Kwon, Park, Jeon, Weon, 2015; List, Gulay, Stoop, Lorenzetti, 2013; Macrum, Bell, Boling, Lewek, Padua, 2012). Attempts to monitor squat depth in sport science research have included practitioner observation, physical aids (e.g. bands, goniometers), and video analysis. However, the subjective nature of practitioner observation subjects this method to inter-rater variability, whereas physical aids can be challenging to replicate between studies. Further, the incorrect use of goniometers can affect its accuracy with respect to the location of bony landmarks, the estimation of the centre of rotation of the joint and the ability to locate and maintain the centre of the goniometer over this point (Gajdosik, Bohannon, 1987). Consequently, 3-dimensional (3D) motion-capture systems are relied upon as the "gold standard" to provide reliable and valid objective feedback. Nonetheless, the accuracy of 3D motion-capture systems comes at the extensive cost of time and resources which many practitioners do not possess.

With this background, cost effective 2-dimensional (2D) motion analysis systems are becoming an increasingly viable option in quantifying lower extremity kinematics (Olson, Chebny, Willson, Kernozek, Straker, 2011). While a plethora of 2D applications have been validated in physical therapy and clinical domains, most of the literature has investigated single joint movements or screening exercises (Keogh et al., 2019). The Coach's Eye is an affordable smartphone 2D motion-capture tool capable of providing joint kinematic feedback from a wide range of movement tasks via its touchscreen goniometer application. The Coach's Eye may provide useful objective feedback through the analysis of peak flexion angles at the hip, knee, and ankle joints. Surprisingly, while the Coach's Eye has been downloaded more than one million times (Mousavi et al., 2020), no study has examined all facets of the application's reliability.

Previous examinations of the Coach's Eye have displayed encouraging validity and reliability findings during treadmill running and wheelchair propulsion (Alkhateeb, Forrester, Daher, Martin, Alonazi, 2017; Mousavi et al., 2020). Though the relevance of these studies to complex movements such as the back squat are limited. In 2015, Krause et al. investigated the test re-test reliability and validity of kinematics during an unloaded squat pattern using the Coach's Eye against a 3D motion-capture system. Acceptable test re-test reliability at the hip (intraclass correlation coefficient [ICC] = 0.98), knee (ICC = 0.98), and ankle (ICC = 0.79) was reported. However, while the reporting of relative reliability statistics (i.e. ICC, r) is undoubtedly of importance, we wish to highlight a series of limitations. One, the omission of a paired samples t test and assessment of measurement error (i.e. coefficient of variation [CV]) prevents any worthwhile conclusions regarding the applications ability to detect meaningful change which isn't the result of measurement error. Another key absence is that of intrarater reliability analysis, which quantifies a single practitioner's self-consistency in scoring (Gwet, 2008). It is of material importance this is quantified because the accuracy of the Coach's Eye depends on the ability of the user to select specific video frames and to draw joint angles via touchscreen (Keogh et al., 2019; Mills, 2015). Moreover, the application's interrater reliability, defined as the agreement between multiple examiners, is not yet known (Koo, Li, 2016).

Together, the issues of intrarater and interrater reliability of the Coach's Eye are imperative because coaches and clinician's decisions are often based on repeated measures by the same or by different examiners. Interestingly, other smartphone goniometer applications have displayed high intrarater and interrater reliability between experienced and inexperienced practitioners (Mehta, Bremer, Cyrus, Milligan, Oliashirazi, 2021; Milanese et al., 2014; Svensson et al., 2019), but it is inadvisable that the findings from one goniometer application should be used to infer the reliability of another. Given the aforementioned widespread use of the Coach's Eye it is reasonable to assume the application is being used by a population with a wide variety of kinematic knowledge; ranging from novice users to experienced users. Consequently, it is of material importance the interrater reliability between novice and expert users is assessed. No study has assessed the test re-test reliability, intrarater and interrater reliability of the Coach's Eye during back squat exercise. This warrants further investigation.

The primary objective of this study was to investigate the test re-test reliability of peak flexion angles of the hip, knee, and ankle joints from the Coach's Eye during back squat exercise. The secondary objective was to determine the intrarater reliability of measures using the same examiner, and the interrater reliability of measures between an experienced and inexperienced examiner. It was hypothesised the test re-test reliability, intrarater reliability, and interrater reliability of the Coach's Eye would be very high.

Methodology

Design

A repeated-measures within-subject design investigated the reliability of joint kinematics during the free-weight back squat. Each participant's back squat 1-repetition maximum (1-RM) was assessed, followed by 2 identical trials utilizing incremental loads of 20%, 40%, 60%, 80%, and 90% 1-RM. Participants were allowed to use their own lifting footwear.

Examiners

The first rater was the primary researcher who had 6 years' applied experience as a sports medicine practitioner. The second examiner was a postgraduate student with less than 1 years' applied experience as a sport scientist. Both examiners underwent a standardization session to familiarise themselves with the data collection methods prior to the study's commencement. Both examiners were blind to the other rater's measurements until all the data had been analysed.

Subjects

A total of 22 strength-trained male weightlifters (mean ± SD; age = 25.0 ±2.6 y; body mass = 90.7 ±14.0 kg; stature = 178.9 ±10.0 cm; back squat = 1-RM 175.7 ±29.2 kg; relative 1-RM = 2.0 ±0.4 x/body mass) were recruited for this study. All subjects had a minimum of 4 years' experience of resistance training and trained approximately 10.1 ±2.7 h per week. A sample size calculation was estimated using G*Power software (Version 3.1.9.3) (Faul, Erdfelder, Lang, Buchner, 2007). To the authors knowledge, no previous estimates of effect size (*ES*) have been established for the Coach's Eye. Twenty-two subjects were required to identify differences between 2 dependant means using a Cohen's d_z of 0.63 (moderate effect), a 2-sided α level of 0.05 and a 1 – β of 0.80. Informed consent was provided prior to data collection with ethical approval granted by the St Mary's University ethics committee in accordance with the seventh revision of the Declaration of Helsinki (2013).

Facilities

Humidity (%) and temperature (°C) were monitored (Govee Thermometer Hygrometer H5075; Govee RGBIC, Los Angeles, CA). All sessions were performed at a similar time of day (±1 h) and were separated by 48–72 h. Subjects were instructed to refrain from strenuous exercise, and to avoid alcohol and caffeine consumption within 24 h of testing throughout the study duration.

Maximum strength assessment

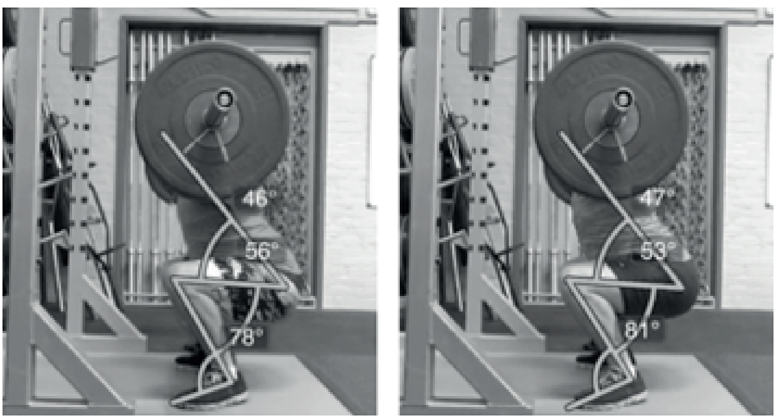
The aims of the first session were to collect subject's anthropometric measures and to assess back squat 1-RM. Body mass (Seca 875; Seca GmbH & Co, Hamburg, Germany) and stature (Seca 202, Seca GmbH & Co, Hamburg, Germany) were recorded. Subjects performed a standardised warm-up protocol, which was used for all sessions. The warm-up consisted of 5 minutes cycling at 60 RPM and 60 W using an air-braked cycle ergometer (Wattbike Pro, Wattbike Ltd, Nottingham, UK) followed by 5 mobility exercises and 10 repetitions with an unloaded barbell. All repetitions were performed using a squat stand or power cage (Eleiko®, Halmstad, Sweden) in conjunction with a calibrated 20 kg barbell and bumper plates (Eleiko®, Halmstad, Sweden). The National Strength and Conditioning Association (NSCA) guidelines for assessing back squat 1-RM were adhered to (Haff, Triplett, 2016). Participants completed 5 repetitions at 50% of estimated 1-RM, 3 repetitions at 70% and 80% of estimated 1-RM, and finally, 90% of estimated 1-RM for a single repetition. As participants approached their estimated 1-RM, loads were increased by 1–10 kg in order to find a true 1-RM for each individual. A maximum of 5 1-RM attempts were allowed. If an attempt was unsuccessful, participants were allowed another attempt with a reduced load. Rest periods were 3 minutes between warm-up sets and up to 5 minutes between 1-RM attempts. Adequate squat depth was confirmed using video footage and observation from a sports medicine practitioner with 6 years' experience. Each subject's preferred feet placement was marked on the ground with a marker pen and white tape.

Joint kinematic assessment

Sessions 2 and 3 were identical; each requiring participants to perform 3 repetitions at 20, 40, 60 and 80% 1-RM and 2 repetitions at 90% 1-RM. Up to 3 minutes rest was provided between sets. All relative loads were rounded up to the nearest 1 kg. Participants were instructed to control the eccentric portion of the back squat at a self-selected pace until full knee flexion (>120.0°) was achieved (Bryanton, Kennedy, Carey, Chiu, 2012), followed by execution of the concentric portion until full hip and knee extension was achieved. Only the repetitions with the deepest squat depth at each loaded condition were analysed. Multiple repetitions were performed to ensure maximum depth was achieved.

Data acquisition

All footage was captured via a smartphone camera system (iPhone 11, version iOS 14.4.2; Apple, Cupertino, CA) utilising the Coach's Eye (TechSmith Corporation, USA, version 6.5.3.0) application at 60 fps and resolution of 1080 p. To minimise measurement error (Whiteley, 2015), the smartphone was rigged onto a tripod set at a height of 62 cm (floor to camera) and distance of 250 cm (camera to centre of the lifting area) in the sagittal plane. The camera configuration was performed by the primary researcher throughout the study duration. Using the application's built-in feature, the video frame showing each subject's lowest portion of the squat at each relative intensity from both trials were displayed on the screen simultaneously (Figure 1). All linear angle markings were drawn via the built-in angle tool with the aid of a touch screen stylus (Mousavi et al., 2020). Markings were applied to anatomical regions previously described in the literature (Schurr, Marshall, Resch, Saliba, 2017): hip flexion was measured as the angle between the acromioclavicular joint and lateral knee joint with the greater trochanter serving as the fulcrum. Knee flexion was measured as the angle between the greater trochanter and lateral malleolus with the lateral knee joint serving as the fulcrum. Ankle dorsiflexion was measured as the angle between a line from the lateral knee joint through the lateral malleolus and a line parallel with the fifth metatarsal. To assess intrarater reliability a single practitioner performed the 2D analysis twice separated by a five-day period (Mousavi et al., 2020). While interrater reliability was determined through the comparison of both examiner's kinematic assessments from the first trial (Romero-Franco, Jiménez-Reyes, González-Hernández, Fernández-Domínguez, 2020).



A – trial 1 B – trial 2

Figure 1. Peak flexion angles at the hip, knee, and ankle captured using the Coach's Eye application

Statistical analysis

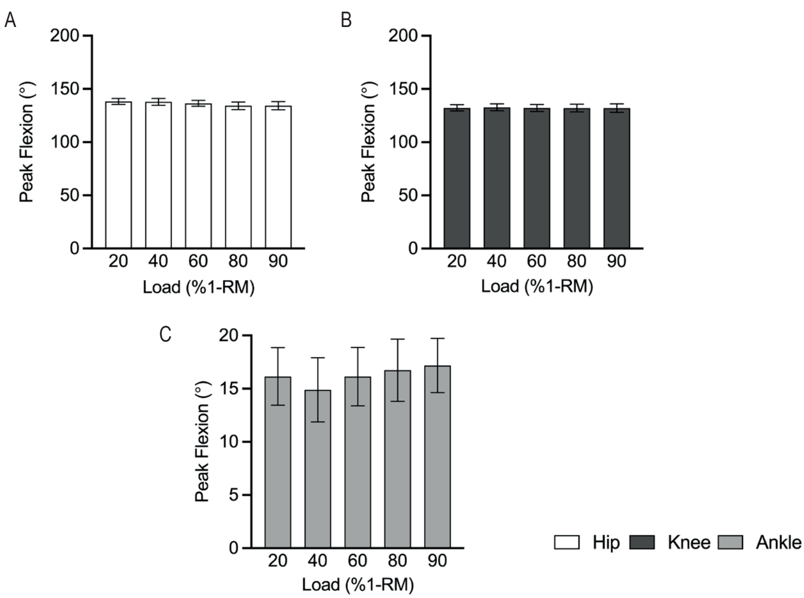
All measures were tested for normality using the Shapiro-Wilk test at an α level of 0.05. All data are presented as mean ± SD unless stated otherwise. Test re-test reliability of outcome measures from Coach's Eye application were assessed at each relative intensity against the magnitude of the correlation coefficient ($ICC_{3,1}$), CV, and *ES*. ICC was also used to determine the intrarater reliability ($ICC_{3,1}$) and interrater reliability ($ICC_{2,1}$) for the kinematic measures from the Coach's Eye (Shrout, Fleiss, 1979). The strength of the correlations were determined using the following criteria: trivial (0.00–0.09), small (0.10–0.29), moderate (0.30–0.49), large (0.50–0.69), very large (0.70–0.89), or nearly perfect (0.90–1.0) (Hopkins, Marshall, Batterham, Hanin, 2009). The magnitude of the CV were categorised as poor (>10%), moderate (5–10%), or good (<5%) (Duthie, Pyne, Hooper, 2003). The magnitude of the *ES* were considered trivial (<0.19), small (0.2–0.59), moderate (0.60–1.19), large (1.20–1.99), or very large

(>2.0) (Hopkins et al., 2009). This study considered the variables highly reliable if they met the following 3 criteria: very large correlation (>0.70) (Lachenbruch, Cohen, 1989), moderate CV (≤10%) (Atkinson, Nevill, 1998), and a small *ES* (<0.60) (Batterham, Hopkins, 2006). The standard error of the measurement (SEM) was also determined (Beckerman et al., 2001; Roebroeck, Harlaar, Lankhorst, 1993), which was used to calculate the minimal detectable change (MDC). The MDC was calculated using the formula (Schmitt, Di Fabio, 2004):

$$MDC = 1.96 \times SD\sqrt{2}\left(-ICC\right)$$

Significant differences of joint angles assessed by the first examiner between both trials were assessed using a 2-tailed paired samples t test with Bonferroni corrections and type 1 error rate set at $\alpha < 0.05$. The significant level was set at $p < 0.05$ and the confidence intervals (CI) for all analyses were set at 95%. The test re-test reliability were performed via a custom spreadsheet (Hopkins, 2015), whereas all other analyses were performed on SPSS (version 27.0: SPSS Inc, Chicago, IL).

Results



A – peak hip flexion; B – peak knee flexion; C – peak ankle flexion.
Error bars indicate SD. 1-RM indicates 1-repetition maximum.

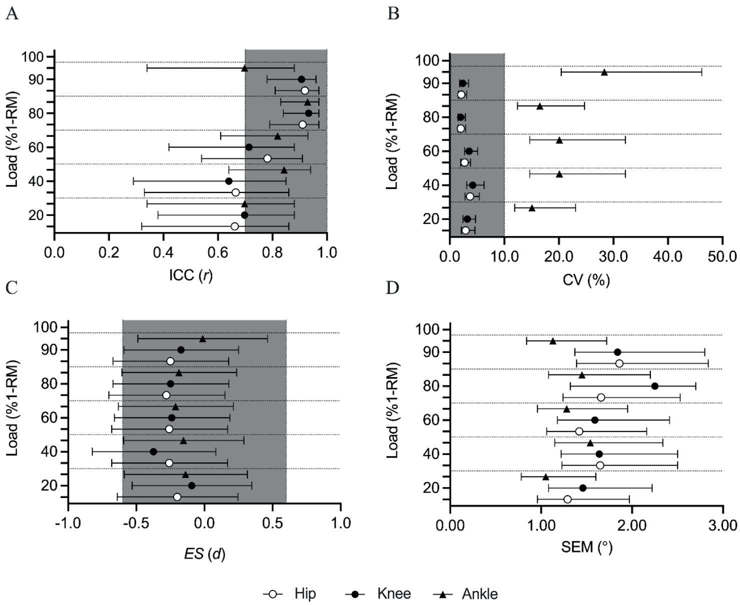
Figure 2. Group mean (SD) values from trials 1 and 2 for peak flexion angles at 20, 40, 60, 80, and 90% 1-RM load

Results from the Shapiro-Wilk test confirmed all measures were normally distributed ($p > 0.05$). No significant differences were found for temperature (trial 1: 14.4 ±3.7°C; trial 2: 14.9 ±4.5°C; $t_{21} = -1.00$, p = 0.33, *ES* = –0.24) and humidity (trial 1: 73.3 ±9.9%; trial 2: 72.5 ±8.9%; $t_{21} = 0.38$, p = 0.71, *ES* = 0.82) between trials. Figure 2 illustrates the overall mean flexion angles assessed by the first examiner. Group means of peak flexion angles between trials are presented in Table 1. No significant differences were detected between trials. The test re-test reliability results of peak flexion angles are shown in figure 3. Peak hip flexion was found to be reliable between 60–90% 1-RM. However, the ICC at 20 and 40% 1-RM did not meet the acceptable reliability threshold. Peak knee flexion was considered reliable at all relative intensities, except for 40% 1-RM, which displayed an ICC < 0.70. Peak ankle flexion was found to be unreliable across all relative intensities. This can be attributed to the poor CV. The intrarater and interrater reliability of peak flexion angles are shown in Table 2. The ICC of peak hip flexion at 20% 1-RM were very largely correlated between rater assessments. All other ICC were deemed to have nearly perfect correlations for peak hip (40–90% 1-RM), knee (20–90% 1-RM) and ankle (20–90% 1-RM) flexion between rater assessments. The interrater agreement displayed nearly perfect correlations across all joints and loaded conditions, with only 2 exceptions: hip flexion at 80% 1-RM and ankle flexion at 20% 1-RM which both showed very large correlations. The MDC of the outcome measures are shown in Table 3.

Table 1. Assessment of significant differences for peak flexion angles at the hip, knee and ankle joints between trials 1 and 2 at each relative intensity using the paired samples t test

Variable	Trial 1	Trial 2	t test[a]	p value
Peak hip flexion, mean ± SD, °				
20% 1-RM	136.6 ±5.4	137.7 ±6.8	–0.89	0.38
40% 1-RM	136.8 ±8.0	139.0 ±8.7	–1.41	0.18
60% 1-RM	136.0 ±6.7	137.2 ±7.5	–1.21	0.24
80% 1-RM	133.5 ±8.6	134.6 ±8.1	–1.31	0.21
90% 1-RM	133.7 ±9.4	134.7 ±9.3	–1.17	0.26
Peak knee flexion, mean ± SD, °				
20% 1-RM	131.0 ±7.3	131.6 ±7.20	–0.42	0.68
40% 1-RM	131.3 ±8.7	134.2 ±7.9	–1.67	0.11
60% 1-RM	131.3 ±8.6	132.8 ±7.2	–1.12	0.27
80% 1-RM	131.2 ±9.4	132.1 ±8.3	–1.16	0.26
90% 1-RM	131.4 ±9.9	132.1 ±8.4	–0.81	0.43
Peak ankle flexion, mean ± SD, °				
20% 1-RM	16.6 ±5.0	17.1 ±5.7	–0.60	0.55
40% 1-RM	14.6 ±8.0	15.2 ±7.8	–0.69	0.50
60% 1-RM	15.8 ±7.1	16.6 ±5.7	–0.99	0.33
80% 1-RM	16.5 ±6.9	17.1 ±7.6	–0.88	0.39
90% 1-RM	18.8 ±4.1	18.8 ±4.8	–0.05	0.96

Abbreviations: 1-RM – 1-repetition maximum.
[a] The degrees of freedom (*df*) = 21, unless otherwise stated.



A – ICC; B – CV; C – ES; D – SEM.
Gray-shaded area indicates the zone of acceptable reliability.
Error bars indicate 95% confidence limits. 1-RM indicates 1-repetition maximum; ICC – intraclass correlation coefficient; CV – coefficient of variation; ES – effect size; SEM – standard error of the measurement.

Figure 3. Forest plot displaying the test re-test reliability of peak flexion angles of the hip, knee, and ankle during the back squat at 20, 40, 60, 80, and 90% 1-RM load

Table 2. Intrarater and interrater reliability of joint kinematics[a]

Variable	Intrarater reliability		Interrater reliability[c]
	Trial 1	Trial 2	
	ICC[b] (95% CI)	ICC (95% CI)	ICC (95% CI)
1	2	3	4
Peak hip flexion °			
20% 1-RM	0.93 (0.82–0.96)†	0.94 (0.86–0.98)†	0.94 (0.84–0.98)†
40% 1-RM	0.91 (0.80–0.96)†	0.93 (0.83–0.97)†	0.94 (0.84–0.98)†
60% 1-RM	0.94 (0.85–0.99)†	0.93 (0.83–0.97)†	0.93 (0.83–0.97)†
80% 1-RM	0.97 (0.89–0.99)†	0.91 (0.80–0.96)†	0.79 (0.53–0.91)†
90% 1-RM	0.96 (0.90–0.98)†	0.95 (0.93–0.99)†	0.95 (0.87–0.98)†
Peak knee flexion °			
20% 1-RM	0.96 (0.89–0.98)†	0.93 (0.83–0.97)†	0.92 (0.80–0.097)†
40% 1-RM	0.97 (0.93–0.99)†	0.96 (0.90–0.98)†	0.96 (0.89–0.99)†
60% 1-RM	0.97 (0.93–0.99)†	0.96 (0.90–0.98)†	0.96 (0.89–0.098)†
80% 1-RM	0.97 (0.96–0.99)†	0.96 (0.90–0.98)†	0.98 (0.94–0.99)†
90% 1-RM	0.98 (0.96–1.00)†	0.95 (0.89–0.89)†	0.99 (0.98–1.00)†

1	2	3	4
Peak ankle flexion °			
20% 1-RM	0.85 (0.65–0.94)†	0.85 (0.66–0.94)†	0.77 (0.48–0.91)†
40% 1-RM	0.87 (0.72–0.95)†	0.92 (0.82–0.97)†	0.92 (0.80–0.97)†
60% 1-RM	0.97 (0.93–0.97)†	0.89 (0.76–0.95)†	0.92 (0.81–0.97)†
80% 1-RM	0.96 (0.92–0.99)†	0.90 (0.77–0.96)†	0.96 (0.90–0.98)†
90% 1-RM	0.94 (0.85–0.98)†	0.93 (0.83–0.97)†	0.91 (0.77–0.96)†

Abbreviation: ICC, intraclass correlation coefficient; CI, confidence interval.
[a] Analyses were restricted to participants without missing values.
[b] ICC are reported as mean at a 95% confidence interval.
[c] Interrater reliability assessed measurements between raters from trial 2.
† p values are significant at <0.001.

Table 3. Recommendations for the minimal detectable change of peak flexion angles at 20, 40, 60, 80 and 90% 1-RM

Load (% 1-RM)	Peak Hip Flexion °	Peak Knee Flexion °	Peak Ankle Flexion °
20	3.6	4.0	2.9[a]
40	4.6	4.5	4.3[a]
60	3.9	4.4	3.6[a]
80	4.6	4.9	4.0[a]
90	5.2	5.1	3.1[a]

Abbreviation: 1-RM – 1-repetition maximum; CV – coefficient of variation; *ES* – effect size; ICC – intraclass correlation coefficient.
[a] Did not meet reliability criteria (ICC > 0.70; CV ≤ 10% and *ES* < 0.60).

Discussion

This was the first study to assess the test re-test, intrarater and interrater reliability of peak flexion angles from the Coach's Eye during back-squat exercise. The primary findings affirm peak hip and knee flexion were reliable across 20–90% 1-RM, while peak ankle flexion was not reliable under any loaded condition. The secondary findings infer the Coach's Eye can produce repeatable assessments of joint kinematics using either a single examiner or 2 examiners, regardless of one's experience.

Joint kinematics remained stable across all loaded conditions. Of relevance, >120.0° of knee flexion was observed at each relative intensity, demonstrating a deep squat depth was achieved (Bryanton et al., 2012). Although supportive literature is limited, 1 study found peak flexion angles (hip = 127.2 ±15.5°; knee = 114.9 ±15.9°; ankle = 27.2 ±5.3°) captured through Coach's Eye are comparable to a 3D motion-capture system (Krause et al., 2015). Bland Altman analysis revealed large systematic bias at the hip (39.8° [–10.3° to –69.3°]), but acceptable bias at the knee (5.0° [–17.6° to 7.6°]), and ankle (3.1° [–14.6° to 8.3°]). Over estimations of hip range of motion highlight a limitation of 2D motion capture systems. This stems from the Coach's Eye's use of linear markers which are unable to account for lumbar-sacrum flexion around the pelvis (Norkin, White, 2009). Practitioners seeking to prioritise lumbar-sacrum assessments are advised to consider 3D kinematic tools (Chowdhury, Byrne, Zhou, Zhang, 2018; Eltoukhy et al., 2016). That aside, very large ICC between trials (Hip: ICC = 0.98; knee: ICC = 0.98; ankle: ICC = 0.79) were found, which coincide with our results. A novel discovery, however, was the high variation observed at the ankle joint across all loaded conditions. This may be explained by inter individual variances in ankle dorsiflexion range of motion (Macrum et al., 2012), or type of footwear worn (Legg, Glaister, Cleather, Goodwin, 2017; Sinclair, McCarthy, Bentley, Hurst, Atkins, 2015). Regrettably, these were not accounted for. High variation may also be explained by the application of linear angles onto anatomical regions without the assistance of reflective markers. Although the absence of markers may be considered a time efficient advantage, this likely reduced the repeatability of measurements. For instance, previous investigations of alternate 2D kinematic systems have shown the assessment of ankle flexion is prone to more error than other joints (Mohammad, Elattar, Elsais, AlDajah, 2021; Romero-Franco et al., 2020). Nevertheless this study's excellent intrarater reliability suggests that joint kinematics are highly consistent when assessed by a single examiner, including at the ankle joint.

This study's intrarater reliability results concur with lower body assessments in the sagittal plane with comparable 2D motion-capture systems (Damsted, Nielsen, Larsen, 2015; Pipkin, Kotecki, Hetzel, Heiderscheit, 2016; Rabin, Einstein, Kozol, 2018). Similarly, our favourable interrater reliability findings are also concurrent with the literature (Mehta et al., 2021; Milanese et al., 2014; Svensson et al., 2019). An intriguing discovery, however, was the relatively lower ICC for ankle flexion at 20% 1-RM. While still acceptable, this too has been observed by Vohralik, Bowen, Burns, Hiller, and Nightingale (2015). It appears the literature's inconsistent reliability results for ankle flexion may simply reflect the lack of agreement between the examiners, rather than the (im)precision of a given goniometer application. In this regard, the Coach's Eye may share the same limitation as the standard goniometer in terms of the subjectivity of establishing body landmarks (Gajdosik, Bohannon, 1987). Nonetheless, this study found an inexperienced and experienced S&C coach can determine joint kinematics with very high agreement. Practitioners should be cognisant of the benefits and limitations of different goniometer applications and how this relates to their place of practice.

A curious finding was the low ICC for peak flexion at the hip and knee joint between trials at 20–40% of 1-RM. This can be explained by the homogeneity of the data observations between trials, which often displayed the exact same values. Such low variability within a sample is known to skew ICC variables (Koo, Li, 2016). This exposes the limitations of relying on a single metric for reliability analysis. Considering the trivial to small *ES* and good CV, peak hip and knee flexion can be considered to have acceptable reliability across 20–90% 1-RM. The MDC reported herein are a slight improvement on values reported by Krause et al. (2015). This may be explained by the video capture speed (60 fps) used in this study. Previous investigations captured footage at 30 fps which causes image blurring (Mills, 2015), and contributes to measurement error (Sheerin, Kendall, Ferber, 2009). Concludingly, considering changes in knee range of motion contribute most to squat depth in the sagittal plane ($r = 0.92$; $p < 0.001$) (Zawadka, Smolka, Skublewska-Paszkowska, Lukasik, Gawda 2020), peak knee flexion from the Coach's Eye may be used to assess squat depth. Given that knee range of motion assessment is prevalent in therapeutic literature (Milanese et al., 2014), the Coach's Eye may be useful in clinical practice. Future research may wish to assess the feasibility of the Coach's Eye, or similar goniometer applications (Weiler, 2016; Vercelli, Sartorio, Bravini, Ferriero, 2017), against 3D kinematic systems using a wider range of rehabilitation exercises (Comfort, Jones, Smith, Herrington, 2015).

Practical implications

The present study shows that peak knee flexion from the Coach's Eye can be used to accurately monitor squat depth using 2 examiners, regardless of experience. To ensure consistency, the equipment setup must be identical between sessions. Further, to aid the validity of longitudinal monitoring the same app and camera system should be used where possible. Because these findings are limited to healthy individuals with no pathologies further research is required to determine whether the Coach's Eye's is a feasible clinical tool for physical therapists. Finally, future studies may also with to determine the validity and reliability of the Coach's Eye during single leg screening exercises or dynamic range of motion tasks (Keogh et al., 2019).

Conclusions

The present study elucidates the Coach's Eye can be used to monitor squat depth in the sagittal plane using multiple examiners with different levels of experience in the full depth back squat using strength-trained males. Caution is advised when using goniometer applications to assess ankle range of motion.

Acknowledgements

The authors would like to thank all of the participants who volunteered for this study. The findings from this investigation do not constitute endorsement of any products assessed by the authors or the journals.

References

Alkhateeb, A.M., Forrester, B.J., Daher, N.S., Martin, B.D., Alonazi, A.A. (2017). Validity and reliability of wheelchair sitting posture measures using Coach's Eye in abled subjects. *Null*, *29* (4), 210–216. DOI: 10.1080/10400435.2016.1220994.

Atkinson, G., Nevill, A. M. (1998). *Statistical Methods For Assessing Measurement Error (Reliability) in Variables Relevant to Sports Medicine*. Adis International. DOI: 10.2165/00007256-199826040-00002.

Batterham, A.M., Hopkins, W.G. (2006). Making meaningful inferences about magnitudes. *International Journal of Sports Physiology and Performance*, *1* (1), 50–57. DOI: 10.1123/ijspp.1.1.50.

Beckerman, H., Roebroeck, M.E., Lankhorst, G.J., Becher, J.G., Bezemer, P.D., Verbeek, A.L.M. (2001). Smallest Real Difference, a Link between Reproducibility and Responsiveness. *Quality of Life Research*, *10* (7), 571–578. DOI: 10.1023/A:1013138911638.

Bryanton, M.A., Kennedy, M.D., Carey, J.P., Chiu, L.Z.F. (2012). Effect of Squat Depth and Barbell Load on Relative Muscular Effort in Squatting. *Journal of Strength and Conditioning Research*, *26* (10), 2820–2828. DOI: 10.1519/JSC.0b013e31826791a7.

Chowdhury, S.K., Byrne, R.M., Zhou, Y., Zhang, X. (2018). Lumbar Facet Joint Kinematics and Load Effects During Dynamic Lifting. *Human Factors*, *60* (8), 1130–1145. DOI: 10.1177/0018720818790719.

Comfort, P., Jones, P.A., Smith, L.C., Herrington, L. (2015). Joint Kinetics and Kinematics During Common Lower Limb Rehabilitation Exercises. *Journal of Athletic Training*, *50* (10), 1011–1018. DOI: 10.4085/1062-6050-50.9.05.

Damsted, C., Nielsen, R.O., Larsen, L.H. (2015). Reliability of video-based quantification of the knee- and hip angle at foot strike during running. *International Journal of Sports Physical Therapy*, *10* (2), 147–154.

Duthie, G., Pyne, D., Hooper, S. (2003). The reliability of video based time motion analysis. *Journal of Human Movement Studies*, *44*, 259–272.

Eltoukhy, M., Travascio, F., Asfour, S., Elmasry, S., Heredia-Vargas, H., Signorile, J. (2016). Examination of a lumbar spine biomechanical model for assessing axial compression, shear, and bending moment using selected Olympic lifts. *Journal of Orthopaedics*, *13* (3), 210–219. DOI: 10.1016/j.jor.2015.04.002.

Escamilla, R.F., Fleisig, G.S., Zheng, N., Barrentine, S.W., Wilk, K.E., Andrews, J.R. (1998). Biomechanics of the knee during closed kinetic chain and open kinetic chain exercises. *Medicine and Science in Sports and Exercise*, *30* (4), 556–569. DOI: 10.1097/00005768-199804000-00014.

Faul, F., Erdfelder, E., Lang, A., Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, *39* (2), 175–191. DOI: 10.3758/BF03193146.

Gajdosik, R.L., Bohannon, R.W. (1987). Clinical measurement of range of motion. Review of goniometry emphasizing reliability and validity. *Physical Therapy*, *67* (12), 1867–1872. DOI: 10.1093/ptj/67.12.1867.

Gwet, K.L. (2008). Computing inter-rater reliability and its variance in the presence of high agreement. *British Journal of Mathematical & Statistical Psychology*, *61* (1), 29–48. DOI: 10.1348/000711006X126600.

Haff, G., Triplett, N.T., National Strength & Conditioning Association (2016). *Essentials of strength training and conditioning* (Fourth ed.). Human Kinetics.

Hartmann, H., Wirth, K., Klusemann, M., Dalic, J., Matuschek, C., Schmidtbleicher, D. (2012). Influence of Squatting Depth on Jumping Performance. *Journal of Strength and Conditioning Research*, *26* (12), 3243–3261. DOI: 10.1519/JSC.0b013e31824ede62.

Hopkins, W. (2015). Spreadsheets for analysis of validity and reliability by linear regression. *Sportscience*, *19*, 36–42.

Hopkins, W.G., Marshall, S.W., Batterham, A.M., Hanin, J. (2009). Progressive Statistics for Studies in Sports Medicine and Exercise Science. *Medicine and Science in Sports and Exercise*, *41* (1), 3–12. DOI: 10.1249/MSS.0b013e31818cb278.

Keogh, J.W.L., Cox, A., Anderson, S., Liew, B., Olsen, A., Schram, B., Furness, J. (2019). Reliability and validity of clinically accessible smartphone applications to measure joint range of motion: A systematic review. *PloS One*, *14* (5), e0215806. DOI: 10.1371/journal.pone.0215806.

Kim, S., Kwon, O., Park, K., Jeon, I., Weon, J. (2015). Lower Extremity Strength and the Range of Motion in Relation to Squat Depth. *Journal of Human Kinetics*, *45* (1), 59–69. DOI: 10.1515/hukin-2015-0007.

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. DOI: 10.1016/j.jcm.2016.02.012.

Krause, D.A., Boyd, M.S., Hager, A.N., Smoyer, E.C., Thompson, A.T., Hollman, J.H. (2015). Reliability and accuracy of a goniometer mobile device application for video measurement of the functional movement screen deep squat test. *International Journal of Sports Physical Therapy*, *10* (1), 37–44.

Lachenbruch, P.A., Cohen, J. (1989). *Statistical Power Analysis for the Behavioral Sciences (2nd ed.)*. American Statistical Association. DOI: 10.2307/2290095.

Legg, H.S., Glaister, M., Cleather, D.J., Goodwin, J.E. (2017). The effect of weightlifting shoes on the kinetics and kinematics of the back squat. *Journal of Sports Sciences*, *35* (5), 508–515. DOI: 10.1080/02640414.2016.1175652.

List, R., Gulay, T., Stoop, M., Lorenzetti, S. (2013). Kinematics of the trunk and the lower extremities during restricted and unrestricted squats. *Journal of Strength and Conditioning Research*, *27* (6), 1529–1538. DOI: 10.1519/JSC.0b013e3182736034.

Macrum, E., Bell, D.R., Boling, M., Lewek, M., Padua, D. (2012). Effect of limiting ankle-dorsiflexion range of motion on lower extremity kinematics and muscle-activation patterns during a squat. *Journal of Sport Rehabilitation*, *21* (2), 144–150.

Martinez-Cava, A., Moran-Navarro, R., Sanchez-Medina, L., Gonzalez-Badillo, J.J., Pallares, J.G. (2019). Velocity- and power-load relationships in the half, parallel and full back squat. *Journal of Sports Sciences*, *37* (10), 1088–1096. DOI: 10.1080/02640414.2018.1544187.

Mehta, S.P., Bremer, H., Cyrus, H., Milligan, A., Oliashirazi, A. (2021). Smartphone goniometer has excellent reliability between novice and experienced physical therapists in assessing knee range of motion. *Journal of Bodywork and Movement Therapies*, *25*, 67–74. DOI: 10.1016/j.jbmt.2020.11.021.

Milanese, S., Gordon, S., Buettner, P., Flavell, C., Ruston, S., Coe, D., O'Sullivan, W., McCormack, S. (2014). Reliability and concurrent validity of knee angle measurement: Smart phone app versus universal goniometer used by experienced and novice clinicians. *Manual Therapy*, *19* (6), 569–574. DOI: 10.1016/j.math.2014.05.009.

Mills, K. (2015). Motion analysis in the clinic: There's an app for that. *Journal of Physiotherapy*, *61* (1), 49–50. DOI: 10.1016/j.jphys.2014.11.014.

Mohammad, W.S., Elattar, F.F., Elsais, W.M., AlDajah, S.O. (2021). Validity and Reliability of a Smartphone and Digital Inclinometer in Measuring the Lower Extremity Joints Range of Motion. *Montenegrin Journal of Sports Science and Medicine*, *10* (2), 47–52. DOI: 10.26773/mjssm.210907.

Mousavi, S.H., Hijmans, J.M., Moeini, F., Rajabi, R., Ferber, R., van der Worp, H., Zwerver, J. (2020). Validity and reliability of a smartphone motion analysis app for lower limb kinematics during treadmill running. *Physical Therapy in Sport: Official Journal of the Association of Chartered Physiotherapists in Sports Medicine*, *43*, 27–35. DOI: 10.1016/j.ptsp.2020.02.003.

Myer, G.D., Paterno, M.V., Ford, K.R., Hewett, T.E. (2008). Neuromuscular training techniques to target deficits before return to sport after anterior cruciate ligament reconstruction. *Journal of Strength and Conditioning Research*, *22* (3), 987–1014. DOI: 10.1519/JSC.0b013e31816a86cd.

Norkin, C.C., White, D.J. (2009). *Measurement of joint motion; a guide to goniometry*, *4th ed*. Ringgold, Inc.

Olson, T.J., Chebny, C., Willson, J.D., Kernozek, T.W., Straker, J.S. (2011). Comparison of 2D and 3D kinematic changes during a single leg step down following neuromuscular training. *Physical Therapy in Sport*, *12* (2), 93–99. DOI: 10.1016/j.ptsp.2010.10.002.

Pipkin, A., Kotecki, K., Hetzel, S., Heiderscheit, B. (2016). Reliability of a Qualitative Video Analysis for Running. *The Journal of Orthopaedic and Sports Physical Therapy*, *46* (7), 556–561. DOI: 10.2519/jospt.2016.6280.

Rabin, A., Einstein, O., Kozol, Z. (2018). Agreement Between Visual Assessment and 2-Dimensional Analysis During Jump Landing Among Healthy Female Athletes. *Journal of Athletic Training*, *53* (4), 386–394. DOI: 10.4085/1062-6050-237-16.

Rhea, M.R., Kenn, J.G., Peterson, M.D., Massey, D., Simão, R., Marin, P.J., Favero, M., Cardozo, D., Krein, D. (2016). Joint-Angle Specific Strength Adaptations Influence Improvements in Power in Highly Trained Athletes. *Human Movement*, *17* (1), 43–49. DOI: 10.1515/humo-2016-0006.

Roebroeck M.E, Harlaar J, Lankhorst G.J. (1993). The application of generalizability theory to reliability assessment: an illustration using isometric force measurements. *Physical Therapy*, *73* (6), 386–401.

Romero-Franco, N., Jiménez-Reyes, P., González-Hernández, J.M., Fernández-Domínguez, J. C. (2020). Assessing the concurrent validity and reliability of an iPhone application for the measurement of range of motion and joint position sense in knee and ankle joints of young adults. *Physical Therapy in Sport*, *44*, 136–142. DOI: 10.1016/j.ptsp.2020.05.003.

Schmitt, J.S., Di Fabio, R.P. (2004). Reliable change and minimum important difference (MID) proportions facilitated group responsiveness comparisons using individual threshold criteria. *Journal of Clinical Epidemiology*, *57* (10), DOI: 1008–1018. S0895435604000666.

Schoenfeld, B.J. (2010). Squatting Kinematics and Kinetics and Their Application to Exercise Performance. *Journal of Strength and Conditioning Research*, *24* (12), 3497–3506. DOI: 10.1519/JSC.0b013e3181bac2d7.

Schurr, S.A., Marshall, A.N., Resch, J.E., Saliba, S.A. (2017). Two-Dimensional Video Analysis is Comparable to 3d Motion Capture in Lower Extremity Movement Assessment. *International Journal of Sports Physical Therapy*, *12* (2), 163–172.

Sheerin, K., Kendall, K.D., Ferber, R. (2009). Measurement error of rearfoot kinematics during running between a 100Hz and 30Hz camera: original research article. *International Sportmed Journal for FIMS*, *10* (3), 152–162.

Shrout, P.E., Fleiss, J.L. (1979). Intraclass correlations: uses in assessing rater reliability. *Psychological Bulletin*, *86* (2), 420–428. DOI: 10.1037//0033-2909.86.2.420.

Sinclair, J., McCarthy, D., Bentley, I., Hurst, H.T., Atkins, S. (2015). The influence of different footwear on 3-D kinematics and muscle activation during the barbell back squat in males. *European Journal of Sport Science*, *15* (7), 583–590. DOI: 10.1080/17461391.2014.965752.

Svensson, M., Lind, V., Löfgren Harringe, M. (2019). Measurement of knee joint range of motion with a digital goniometer: A reliability study. *Physiotherapy Research International: The Journal for Researchers and Clinicians in Physical Therapy*, *24* (2), e1765-n/a. DOI: 10.1002/pri.1765.

Vohralik, S.L., Bowen, A.R., Burns, J., Hiller, C.E., Nightingale, E.J. (2015). Reliability and validity of a smartphone app to measure joint range. *American Journal of Physical Medicine & Rehabilitation*, *94* (4), 325–330. DOI: 10.1097/PHM.0000000000000221.

Weiler, R. (2016). Übersense coach app for sport medicine? slow motion video analysis (mobile app user guide). *British Journal of Sports Medicine*, *50* (4), 255–256.

Whiteley, R. (2015). Coach's eye. *British Journal of Sports Medicine*, *49* (20), 1349–094656. Epub 2015 Jun 5. DOI: 10.1136/bjsports-2015-094656.

Wirth, K., Hartmann, H., Sander, A., Mickel, C., Szilvas, E., Keiner, M. (2016). The Impact of Back Squat and Leg-Press Exercises on Maximal Strength and Speed-Strength Parameters. *Journal of Strength and Conditioning Research*, *30* (5), 1205–1212. DOI: 10.1519/JSC.0000000000001228.

Vercelli, S., Sartorio, F., Bravini, E., Ferriero, G. (2017). DrGoniometer: A reliable smartphone app for joint angle measurement. *British Journal of Sports Medicine*, *51* (23), 1703–1704.

Zawadka, M., Smolka, J., Skublewska-Paszkowska, M., Lukasik, E., Gawda, P. (2020). How Are Squat Timing and Kinematics in The Sagittal Plane Related to Squat Depth? *Journal of Sports Science & Medicine*, *19* (3), 500–507.

Cite this article as: Kilgallon, J., Cave, J., Cushion, E., Joffe, S., Tallent, J. (2022). Reliability of the Coach's Eye Goniometer Application During Squat Exercise. *Central European Journal of Sport Sciences and Medicine*, *2* (38), 47–59. DOI: 10.18276/cej.2022.2-05.