Journal of Equine Rehabilitation 3 (2025) 100023

Contents lists available at ScienceDirect

Journal of Equine Rehabilitation

journal homepage: www.journals.elsevier.com/journal-of-equine-rehabilitation

The Equine Quality of Movement Score: How reliable is it?

A.G. Bowen [a,*], G. Tabor [b], R. Labens [a], M. Douglas [a], H. Randle [a]

[a] School of Agricultural, Environmental and Veterinary Sciences, Charles Sturt University, 1 Boorooma Street, Wagga Wagga, NSW 2678, Australia
[b] Equestrian Performance Research Centre, Hartpury University, Gloucester, Gloucestershire GL 19 3BE, UK

ARTICLE INFO

Keywords:
Equine physiotherapy
Gait
Movement dysfunction
Outcome measure
Rehabilitation

ABSTRACT

Visually assessing equine quality of movement in the field is a routine part of performance management and rehabilitation procedures. To improve this subjective process, an outcome measure to grade quality of movement has been developed. The Equine Quality of Movement Score (EQoMS) is a battery of 30 standardised in-hand movement tests with a paired word and number grading system and directives to guide scoring based on key features. Six assessors (3 equine physiotherapists and 3 specialist equine veterinarians) scored video footage of 20 horses performing these movement tests. Videos were presented for scoring in a randomised order over two rounds with a 2-week gap in between. Intra-rater reliability was excellent (Spearman's rho = 0.95, $P = <0.001$), and inter-rater reliability using ranked scores was moderately strong (Spearman's rho = 0.66–0.95, $P = <0.001$). Assessors agreed on the relative quality of movement but disagreed on the absolute magnitude of movement quality. The assessors' free-text comments indicated a bias towards making negative comments. Reflecting the EQoMS directives, different features were commented on for different gaits, figures of movement and surfaces. Differences in terminology were found in the comments made by specialist equine veterinarians and equine physiotherapists, with the latter avoiding directly commenting on lameness. The variation in scores and comments emphasises the subjective nature of movement quality. Even when given the same training and directives to guide scoring, individual assessors' weighting of qualitative and quantitative elements was different. Further refinement is required to improve agreement on characteristics of equine movement, and the reliability of the EQoMS.

1. Introduction

Movement related outcome measures are necessary for monitoring changes in a horse's condition not just in research settings but also in the field. Quality of movement is an important indicator, for equine physiotherapists and rehabilitation specialists, encompassing whole body biomechanics, gait pattern, symmetry, motor-control, muscle activity and timing, willingness and behaviours indicating discomfort/comfort [1,2]. However, no outcome measures existed to assess this in horses. Currently, subjective visual assessment is used to characterise movement observed in the field [3,4] and is, therefore, a barrier to evidence-informed practice. While instrumented gait analysis is advancing not all practitioners have access and they have their own limitations such as cost, time, biomechanical and gait-limited assessment [5,6]. Without valid and reliable outcome measures unproven therapies may persist, potentially crowding out effective therapies and presenting a significant welfare concern by prolonging horses' suffering and impacting horse owners economically [7–9]. To address this a series of studies were conducted by Bowen et al. [3,10] to develop a new outcome measure to grade equine quality of movement. The Equine Quality of Movement Score (EQoMS) is a battery of 30 standardised in-hand movement tests with a paired word and number grading system and associated directives based on key features to guide scoring.

Assessing the movement quality exhibited by a horse requires standardised, reliable scoring, that is, the ability to detect change not masked by measurement error or observer bias. Two aspects of reliability need to be considered: intra-rater reliability refers to a clinician's consistency over time, while inter-rater reliability is the agreement between two or more clinicians. Intra-rater reliability is fundamental to detecting change over the course of rehabilitation, while high inter-rater reliability is particularly important during complex cases that require interdisciplinary team input [11].

In contrast to instrumented gait analysis's focus on quantifying movement [12,13] and lameness examination's focus on identifying affected limbs and severity [14–16], quality of movement assessment considers both qualitative and quantitative features of movement [17].

* Corresponding author.
E-mail address: abowen@csu.edu.au (A.G. Bowen).

https://doi.org/10.1016/j.eqre.2025.100023
Received 1 October 2024; Received in revised form 29 January 2025; Accepted 28 February 2025
Available online 4 March 2025

This holistic concept encompasses the ability to perform with ease and efficiency [18], extending beyond clinical issues with variation in non-lame horses due other to factors. Several studies have demonstrated varying agreement between veterinarians when assessing single elements such as equine lameness in straight lines [16,19], on circles [15], fitness to compete [20], back mobility [21], hoof pain [22], facial expressions of pain [23] and ridden horse behaviours [24]. Unlike movement asymmetry, which can be reliably quantified [25], quality of movement is a subjective construct; thus, despite standardising processes and training in scoring, complete agreement is unlikely. In the absence of an established reference standard, the EQoMS is a progression from the current judgement call when visually assessing movement, to grading it with an ordinal scale. This is a starting point for clinical practice and with further refinement may become sufficiently reliable for use in research too.

The primary aim of this study was to determine the intra- and inter-rater reliability of the EQoMS. The secondary aim was to determine if the scores and qualitative comments provided by equine physiotherapy and specialist equine veterinary clinicians differ. Given that specialist equine veterinarians receive substantial training in lameness diagnosis, it is hypothesised their comments may focus more on identifying a lame limb, while equine physiotherapists may focus more on motor control, movement pattern, coordination, or muscle activity.

2. Materials and methods

2.1. Ethical approval

The study protocol was approved by Charles Sturt University's (CSU) Animal Ethics Committee (AEC) (Approval number #A23778; 4th October 2023) and Human Research Ethics Committee (Approval number #H23782; 18th September 2023). Informed consent was obtained from all horse owners and assessors involved in the study.

2.2. Recruitment

2.2.1. Assessors

Six independent assessors, three equine physiotherapists and three specialist equine veterinarians were invited. Assessors had not been involved with any of the previous steps leading to the development of the EQoMS [8]. Specialist equine veterinarians were specifically chosen; being diplomates of the American College of Veterinary Sports Medicine and Rehabilitation with a focus on horses. The equine physiotherapists held human and veterinary physiotherapy qualifications, had been practising for many years and are involved with the education of veterinary physiotherapists.

2.2.2. Horses

Two groups of horses from the Charles Sturt University teaching herd were used along with 7 privately owned horses recruited from students and local adult riding club members responding to an expression of interest call. The horses ranged in age from 3 to 27 years (mean = 11.8, S. D. = 5.9), 7 mares, 18 geldings and approximate height of 14.1hh to 16.3hh, mostly thoroughbreds (n = 20), the remaining breeds included Australian Stock Horse, Welsh x Warmblood, Quarter horse and Standardbred (see supplementary item 1 for table of individual horse features). The CSU horses were unridden and used for teaching horse handling, veterinary procedures and in research, whilst the privately owned horses were primarily used for leisure riding and adult riding club activities.

2.3. Data collection

Twenty-five horses took part in data collection at CSU's Wagga Wagga Equine Centre over a 2-week period during December 2023. The majority of the horse handling was performed by the first author (AB) along with some privately owned horses' owners. Over the course of data collection, research assistants were responsible for filming and some horse handling. Author RL or a resident in equine sports medicine were on-call for veterinary assistance if required. RL provided training to AB and MD in setting up and using an inertial measurement unit (IMU) based gait symmetry system (Lameness Locator, Columbia, Missouri, USA) as objective data were concurrently collected for validity analysis. Due to the seasonal heat filming was undertaken in the mornings only.

Horses were observed walking from the paddock/float for basic soundness and well-being. They were then groomed, hooves picked out, a hoof tester applied to all four hooves, and fly repellent spray applied. The IMU system ear bonnet was fitted under the headcollar, the pastern wrap was applied to the right forelimb, and the tuber sacrale sensor was attached centrally with double-sided tape and butterfly clips.

The horses were led through the battery of in-hand movement tests (see supplementary item 2 for EQoMS test sheet) in this order: walking, trotting 30 m up and back on a firm surface, step back a minimum of four steps repeated twice, hindleg cross over full 360° to the left and the right repeated twice, at least two circuits of 5–10 m diameter circles on the left and right rein at walk and trot. These movements (except hindleg cross-over) were repeated on the soft surface before progressing to lungeing at the walk, trot, and canter, aiming for a circle diameter of 15 m+ , concluding with transitions between halt, walk and trot viewed from behind, side on and in front. An assistant operated the video camera (Sony Handycam HDR CX625, on a 1.5 m tripod), starting and stopping at the beginning and end of each movement test, and Lameness Locator on an iPad for trotting in a straight line (firm and soft surfaces) and lunge circles only. A 30 m section of gravel roadway was used for the firm surface and an adjacent sand jumping arena was used for the soft surface. Jumping wings and poles were used to partially enclose a corner of the jumping arena to create an 18 m diameter space for lungeing. The halter was replaced with a lungeing cavesson and a lungeing rein with stitched markers every metre was used to define circle diameter. A dressage whip was carried during in-hand movement tests and a lunge whip during lungeing based movement tests.

On completion of movement tests, IMU system sensors were removed, horses hosed, given a food treat, and returned to their paddock with their herd-mates or to the owner's float/ trailer. Owners were present at all times for privately owned horses. In appreciation for their time, owners were gifted a $25AUS Amazon (Bellevue, Washington, USA) voucher and provided with summary non-diagnostic information for their horse. Owners were also invited to provide feedback via Survey Monkey (San Mateo, California, USA).

Video footage was edited (trimmed and background audio removed on some clips due to conversations that may influence the assessors) using Movie Maker (version 3.7.5, Redmond, Washington, USA). Footage of five horses was incomplete due to lameness or behavioural issues – these were not included. Video files for the remaining 20 horses were assigned unique codes, and a random draw of numbers from a cup was conducted to determine the order these would be presented to the assessors in rounds 1 and 2. Survey Monkey was set up to record the scores and comments for each horse.

Assessors viewed a 45-minute training program video covering the development of the EQoMS and the terminology used in the directives. There were two practice sessions, using video footage of five horses used in the earlier pilot study reported in Bowen et al. [10]. The first practice was movement-test-by-movement-test, and the second practice was horse-by-horse; these totalled 2.5 hours of footage. On completing the training programme, the assessors were asked to self-assess their competence to use the new outcome measure. During the training extra feedback was provided to answer questions from the assessors, see supplementary item 3. Assessors were gifted a $500AUS Amazon voucher in appreciation for their time (15hrs+).

During January and February 2024, the assessors viewed the video footage which was shared via Microsoft OneDrive (Redmond,

Washington, USA) and entered their scores and comments into Survey Monkey (Round 1). Two weeks later, with the videos presented in a different randomized order the assessors repeated the process (Round 2). Finally, the assessors provided feedback on the training program, the EQoMS form and their thoughts on the clinical application of the EQoMS via email to the primary author. The assessors did not have access to the results of the IMU system data collected, which were analysed later.

2.4. Data analysis

Survey Monkey data were downloaded into Microsoft Excel (v.16.0, Redmond, Washington, USA), screened for missing data and identifiers were removed. Non-parametric tests were used as the data were ordinal. The distribution of horses and assessors scores was described via summed totals, frequency counts of scores, movement test medians and range. Order effect was investigated with Spearman's Rho and Kendall's Tau. Spearman's Rho was used for intra-rater reliability and to test for correlations between scores for different movement tests. Inter-rater agreement was explored with kappa statistics and Spearman's Rho for rankings. Independent Mann-Whitney t-tests were used to compare reliability between specialist equine veterinary and equine physiotherapy assessors. The *P*-value was set as 0.05. Statistical analysis was carried out with Jeffery's Amazing Statistical Program (JASP) (v.0.18.3.0). IMU system data were analysed and used in a separate paper addressing validity.

To analyse the qualitative data, comments were input to NVivo [data management and analysis software, v.12] (Denver, Colorado, USA) and coded by movement test, assessor profession and score (0−5). Word frequency counts and matrix queries were used to compare the types of comments given for different gaits, figures, and surfaces, by different professions and for different scores. MS Excel was used to determine the number and length of comments provided.

3. Results

3.1. Descriptive results

All 6 raters completed both rounds. An error occurred in Survey Monkey, labelling a score of 5 as excellent instead of optimal, but that should not affect the results. The horses' total scores, over both rounds, ranged from 24 to 148 out of 150 possible points with a median total score of 85. The distribution of scores between round 1 and round 2 was very similar (Table 1).

Most movement tests gained a median score of 3 (mild movement dysfunction) except the following: small circle trot right rein on a firm surface (median = 2), canter circle left rein (median = 2.5) and trot to canter transitions left rein (median 2.5). For several movement tests horses received higher scores, reflected in a median score of 4 (indicating good quality of movement); walk in a straight line on a soft surface, canter circle right rein, halt to walk transitions and trot to walk transitions.

Table 1
Distribution of Equine Quality of Movement Score grades between round 1 and round 2. Grades were assigned by 6 assessors to 20 horses performing 30 in-hand movement tests. N.B. Grades 0 to 5. Missing values = assessor missed entering a value in Survey Monkey. Percentage in brackets.

Scores	Round 1	%	Round 2	%	Combined	%
5 optimal	273	7.6	285	7.9	558	7.8
4 good	1075	29.9	1011	28	2086	29
3 mild movement dysfunction	1068	29.7	1052	29.2	2120	29.4
2 moderate movement dysfunction	541	15	568	15.8	1109	15.4
1severe movement dysfunction	355	9.9	407	11.3	762	10.6
0 unable to complete test/ assign score	275	7.6	257	7.1	532	7.4
Missing values	13	0.4	20	0.6	33	0.5
Total	3600		3600		7200	

The range of scores allocated by each assessor is illustrated in Fig. 1. In round 2 the assessors scores showed slightly less dispersion however, assessors maintained their individual positions relative to the group. Assessors 1, 2 and 3, the specialist equine veterinarians, generally awarded higher scores, while assessors 4, 5 and 6, the equine physiotherapists, generally awarded lower scores.

The range of total scores for each horse had a median of 77 points difference (see Fig. 2). With total scores out of 150, the variation in total scores between some assessors for some horses is more than half the marks available. For example, horse 11's total scores ranged from 35 to 132 (97 points difference), with a median of 93.5. For horse 7 the total scores were closer ranging from 26 to 69 with a median of 58, with a difference of 43 points.

There was a weak negative effect of order, i.e., each horse's randomly allocated position within the two assessment rounds Spearman's rho = -0.14, $P = 0.02$ and Kendall's Tau = -0.1, $P = 0.02$ (For scatterplot see Supplementary item 4).

3.2. Intra-rater reliability

The intra-rater reliability for total scores was excellent (Spearman's rho = 0.95, $P = <0.001$) (For scatterplot, see Supplementary item 5a). Individually, assessor reliability varied but was still very high, minimum = 0.79, maximum = 0.97, all *P* values < .001 (see supplementary item 5b). Assessors 2 (Spearman's rho = 0.97) and 3 (Spearman's rho = 0.94) had the highest correlations between rounds 1 and 2, while assessor 6 had the lowest intra-rater reliability (Spearman's rho = 0.79). The specialist equine veterinarians (assessors 1, 2 and 3) were slightly more consistent in their scoring between the two rounds (higher intra-rater reliability values) (0.89 to 0.97) than the equine physiotherapists (assessors 4, 5 and 6) (0.79 to 0.89).

3.3. Inter-rater reliability

There was very poor agreement between all raters for total scores in round 1 (k = 0.001, Fleiss kappa = 0.008, Kripendorf's alpha = 0.04) and in round 2 (k = 0.003, Fleiss kappa = −0.007, Kripendorf's alpha = −0.04). When total scores were ranked from 1st to 20th place and tested for agreement on order between assessors there were moderately strong correlations, ranging from 0.66 to 0.95, all *P* values < .001 (see Table 2).

3.4. Comparison of scores between movement tests

Movement tests of interest were chosen for comparison among gaits, surfaces, and figures (straight line or circle). There was a poor correlation among scores obtained for trotting in a straight line on a firm surface and scores for cantering left and right on the lunge (see Table 3). Whereas there was good correlation between performance on left and right rein circles at walk, trot, and canter. There was fair to moderate correlation between scores for walk and trot in a straight line on a firm surface and a soft surface (see Table 3). There was a moderate correlation between scores for walk on a straight line and on a circle. The correlation for trot on a firm surface was only fair between straight line and circles.

3.5. Trends in assessors' comments

Frequency counts of the top 50 words (and stemmed words) revealed the horse's head was commented on most often (n = 618), with behind 2nd (n = 524) and lame or lameness 3rd (n = 416 times). Bend and stiffness were 15th (n = 236) and 16th (n = 224). Pain/painful at 24th

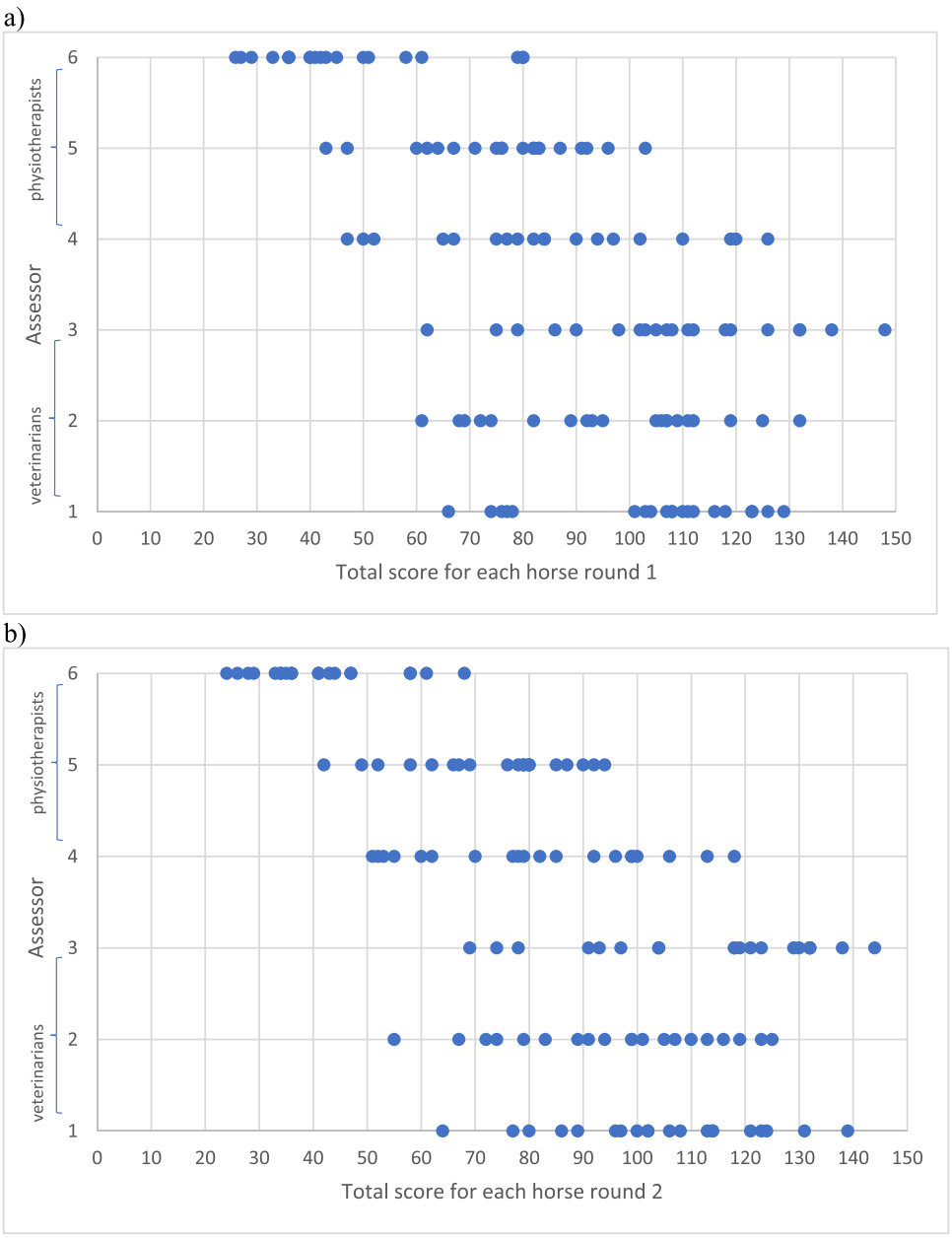

Fig. 1. The range of total scores from each assessor in a) round 1 and b) round 2. There were 20 horses and a maximum score of 150. The higher the score the better the quality of movement. Assessors 1, 2, and 3 were specialist equine veterinarians and assessors 4, 5 and 6 were equine physiotherapists.

(n = 190) followed by weak/weakness in 25th (n = 189). Rhythm was listed at 32nd being stated 147 times. Many of the most frequently used words clearly match those in the directives: Tracking (up or straight), head, tail carriage (tail & swishing), front/behind, spinal, tension, step length (short stepping or striding), rhythm, crosses midline, lateral flexion on circle line (lateral, bend, circle), willing (reluctant), comfort (pain/ face), muscle activity (weakness) un/balance, toe drag (hoof clearance), spinal movement (neck, back) and head position (position). Similarly, several of the most frequently used words are modifiers: lacks, mildly, better, good, still, intermittently, slight, poor.

Some comments occurred more frequently with specific scores, for instance pain, discomfort or lameness were more frequently reported with low scores, while positive descriptors such as good, nice, or lovely were more frequently used with higher scores (see Fig. 3) (see Supplementary item 7 for all frequency counts).

Phrases were also associated with gaits, figures, and surfaces (see Supplementary item 7 for all frequency counts). Some gait specific comments were speed for walk, head nod for trot, and rushed, balance, tense, and wrong lead for canter. Lameness and pain were commented on more often for trot than for walk. Lameness was also commented on more frequently on the circle (n = 270) than the straight line (n = 112). Bend (n = 184) was mentioned for circles but not for straight lines. Step back elicited more comments about willingness, such as reluctance (n = 67) and resistance (n = 15). Lameness (n = 226), pain (n = 83) and short steps (n = 230) were commented on more frequently on a firm surface. While dragging toes (n = 273), weakness (n = 125), hindlimb issues (n = 125) and rhythm (n = 103) were commented more often on the soft surface.

3.6. Comparison between professions

Intra-rater reliability was high for all assessors with no difference between professions (t = -2.33, P = 0.08). There was higher inter-rater reliability among the specialist equine veterinarians (mean = 0.92, S.

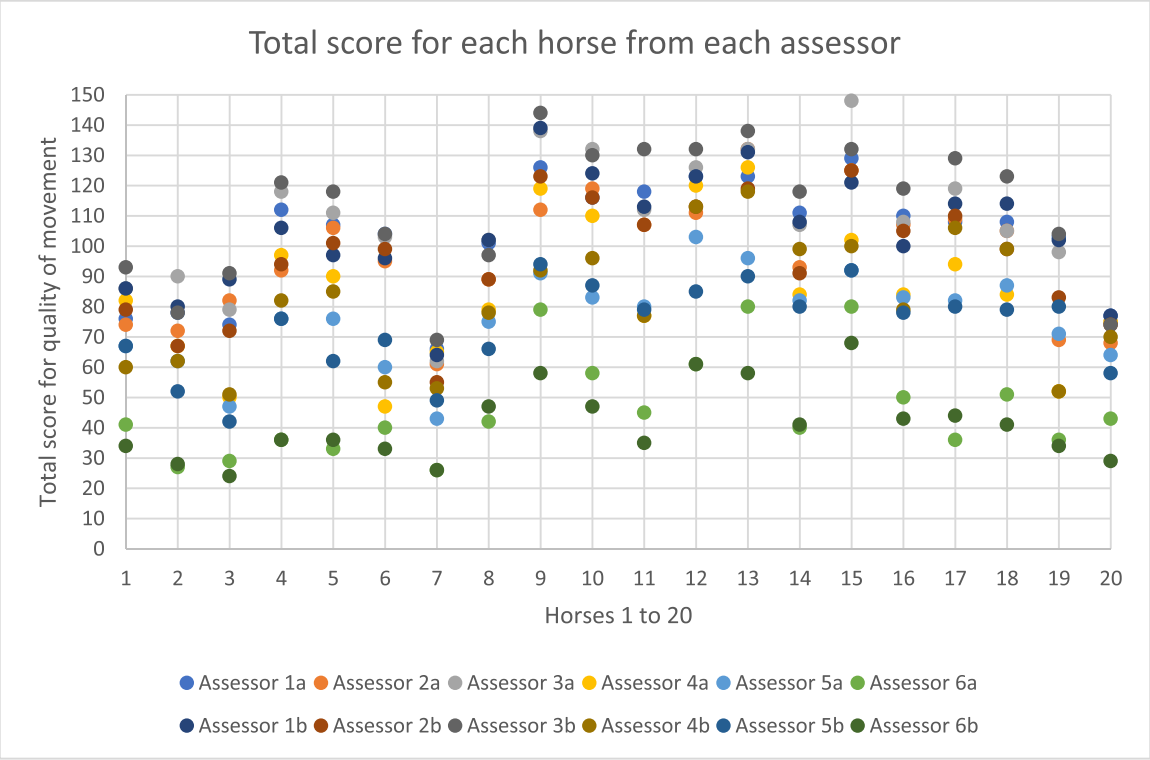

Fig. 2. The range of total scores for each horse from the 6 assessors across two rounds. Each of the 20 horses has 12 total scores displayed. Round 1 and round 2 scores are represented by a and b in the figure legend and different colours.

Table 2
Inter-rater reliability between 6 assessors scoring a group of 20 horses using the Equine Quality of Movement Score over 2 rounds. Individual correlations between assessors when horses total scores are ranked from 1st to 20th a) round 1, b) round 2.

a) **Spearman's Correlations**							
Variable		**Assessor 1a**	**Assessor 2a**	**Assessor 3a**	**Assessor 4a**	**Assessor 5a**	**Assessor 6a**
1. Assessor 1a	Spearman's rho	—					
	p-value	—					
2. Assessor 2a	Spearman's rho	0.88	—				
	p-value	< .001	—				
3. Assessor 3a	Spearman's rho	0.95	0.93	—			
	p-value	< .001	< .001	—			
4. Assessor 4a	Spearman's rho	0.80	0.80	0.86	—		
	p-value	< .001	< .001	< .001	—		
5. Assessor 5a	Spearman's rho	0.89	0.86	0.86	0.88	—	
	p-value	< .001	< .001	< .001	< .001	—	
6. Assessor 6a	Spearman's rho	0.76	0.74	0.67	0.66	0.82	—
	p-value	< .001	< .001	< .001	0.001	< .001	—
b) **Spearman's Correlations**							
Variable		**Assessor 1b**	**Assessor 2b**	**Assessor 3b**	**Assessor 4b**	**Assessor 5b**	**Assessor 6b**
1. Assessor 1b	Spearman's rho	—					
	p-value	—					
2. Assessor 2b	Spearman's rho	0.88	—				
	p-value	< .001	—				
3. Assessor 3b	Spearman's rho	0.94	0.94	—			
	p-value	< .001	< .001	—			
4. Assessor 4b	Spearman's rho	0.80	0.77	0.77	—		
	p-value	< .001	< .001	< .001	—		
5. Assessor 5b	Spearman's rho	0.92	0.86	0.88	0.72	—	
	p-value	< .001	< .001	< .001	< .001	—	
6. Assessor 6b	Spearman's rho	0.86	0.86	0.82	0.85	0.83	—
	p-value	< .001	< .001	< .001	< .001	< .001	—

D. = 0.03) than among the equine physiotherapists (mean = 0.79, S.D. = 0.08) (t Mann-Whitney = 34, P = 0.01). The equine physiotherapists gave lower scores than the specialist equine veterinarians (see Fig. 1); specialist equine veterinarians total scores mean = 103.3, S.D. = 20.8, equine physiotherapists total scores mean = 67.2, S.D. = 24.2, (t = 12,363 P < .001).

The equine physiotherapy participants made more comments than the specialist equine veterinary participants. Out of a total possible 3600 comments, collectively the specialist equine veterinarians in this sample commented 1141 times (32 %) and left 2459 (68 %) comment boxes

Table 3
Correlations between the scores provided by 6 assessors for a group of 20 horses using the Equine Quality of Movement Score for different movement tests.

Movement 1	Movement 2	Spearman's *rho*	P
Trot straight line firm	Canter left lunge soft	0.38	< .001
Trot straight line firm	Canter right lunge soft	0.37	< .001
Canter left lunge soft	Canter right lunge soft	0.90	< .001
Walk straight line firm	Trot straight line firm	0.73	< .001
Walk straight line firm	Walk straight line soft	0.72	< .001
Walk straight line firm	Trot straight line soft	0.61	< .001
Trot straight line firm	Walk straight line soft	0.58	< .001
Trot straight line firm	Trot straight line soft	0.65	< .001
Walk straight line soft	Trot straight line soft	0.67	< .001
Walk straight line firm	Small circle walk left firm	0.76	< .001
Walk straight line firm	Small circle walk right firm	0.73	< .001
Small circle walk left firm	Small circle walk right firm	0.92	< .001
Trot straight line firm	Small circle trot left firm	0.49	< .001
Trot straight line firm	Small circle trot right firm	0.56	< .001
Small circle trot left firm	Small circle trot right firm	0.80	< .001

blank. In contrast, equine physiotherapists in this sample commented 2108 times (59 %) and left 1492 (41 %) comment boxes blank. Additionally, the equine physiotherapists provided longer comments (mean = 10.9 words long), often mentioning multiple features, while specialist equine veterinarians' comments were shorter (mean = 4.5 words long).

Lame/ness was the most frequently used word by specialist equine veterinary assessors (n = 315), with many comments in relation to lameness e.g. rf, lf, rh, lh, (right forelimb, left forelimb, right hindlimb, left hindlimb) bilateral, asymmetry, intermittently. Pain did not occur in their top 50 comments. In contrast, equine physiotherapists used lame/ness a lot less frequently at only 101 times, while pain was commented on 190 times. The specialist equine veterinarians commented on unbalanced 29 times, whereas equine physiotherapists commented on balance 198 times and unbalanced 94 times. Both professions used many of the same words but with different frequencies and emphases.

3.7. Feedback from horse owners and assessors

Four (out of 7) horse owners completed the feedback survey; with 50 % fully and 50 % somewhat understanding the purpose and benefits of the new assessment. All felt that all the movement tests were relevant to their horse, and they felt confident performing all the movement tests. When asked how your horse responded during and after the assessment, there were no adverse events reported; comments were "very well" and "well behaved as always". One owner commented that their horse was lazy during assessment. Overall, anonymous, owners found the process, "Relatively easy for my horse to do," and "There was lots of clear, easy communication in what was being looked for".

All 6 assessors provided feedback. The training program was rated on average 8.8/10 (range 8 - 10) by the assessors. They felt the training was comprehensive, particularly liking the clear explanations and the chance to practice as many times as they wanted. They felt it familiarised them with the process but found it hard not to compare between horses and to change from just assessing a lameness grade. Suggestions to improve the training included: providing written information, more horses, and more information on how to score transitions.

The EQoMS form was rated 7.6/10 on average (range 6 - 8.5). Assessors liked the level of detail, the movements included, being able to review the criteria (directives) easily, and explanations of the parameters of the movements. They described it as "pretty straightforward" and "a useful method for recording to compare to at a later date". However, some assessors suggested reducing the length, adding pictures, providing more space to write comments, and further condensing the directives.

Assessors from both professions liked that the EQoMS considered a lot more aspects of movement than a lameness grading scale and that there was a good array of movements in different conditions included that could inform clinical reasoning. They disliked the length of time it took, some stating there was repetition and suggesting removal of the small circles at walk and trot on the soft surface and going straight to the lungeing and that it was unnecessary to assess step back on both surfaces. Some felt the transitions could also have been assessed on the lunge rather than with separate tests at the end. However, one assessor would have added another test by viewing the hind leg cross-over on the soft surface as well.

Assessors found the level of training and behavioural issues displayed by some horses negatively impacted the scoring process and that some handlers interfered with the movement. Several assessors commented on a lack of range in quality in the horses used meaning they did not get the opportunity to use the higher end of the grading system. Videos were occasionally rewatched if they lost focus. Responses were mixed when asked how likely you would be to use the EQoMS in your practice (6.2/10 on average, range 3 – 10).

4. Discussion

The EQoMSs combination of quantitative scores and qualitative comments is intended to assist in the challenging task of grading the subjective construct of quality of movement in horses. This study confirmed the EQoMS has excellent intra-rater reliability and moderate inter-rater reliability. However, these key measures of reliability are just part of the picture. Variations in scores and comments highlight the subjective nature of the assessment of movement quality and how different aspects are valued.

The assessors were consistent in scoring between rounds 1 and 2, with the distribution of scores between rounds very similar. The weak negative order effect, assessors becoming slightly more critical (less generous) towards the end of a round compared to at the beginning, is counter to the positive order effect reported in dressage competition [26, 27]. The assessors used the full range of scores despite the comments from several that the horses presented had low quality movement. This highlights the circumstance specific nature of quality of movement. The lower scores for trot circles to the right on the firm surface could be explained by horses' reluctance to be led from the right side, many horses broke gait returning to walk, making assessment harder. Most horses in this study were retired racehorses. In this state (New South Wales), thoroughbreds race clockwise (on the right rein) this racing background could be an explanation for higher median scores for the right canter and lower median scores for the left canter [28]. The higher median score for walk on a soft surface, could relate to walk's lower physical demands, with the gait pattern always keeping one hoof on the ground it is lower impact on the limbs and body [29].

Unfortunately, there were some scores of zero, the majority because the horse did not canter, preventing assessment of the canter and transitions, while a few other instances were due to difficulties in performing the required movement. Such as trotting a small circle where poor performance may be attributed to hoof soreness, behavioural or training issues, particularly lack of confidence being handled from the offside (to place the handler on the inside of a circle to the right). Although zero scores reduce a horse's total score, assessor comments can guide reassessment, and individual movement scores can be compared, not just totals.

There was a wide range of total scores. Because quality is a subjective construct it is difficult to know how much of the range in scores is due to differing quality in the horses' movement or due to differing opinions of the assessors. It would be expected that different groups of horses would present a different array of scores. One assessor consistently scored lower than the rest of the group, and their comments reflected this more critical/negative approach. Even with removing assessor 6 the inter-rater reliability was still low. Being subjective there is no right or wrong, perhaps with more discussion during training assessors scores

Fig. 3. Word clouds; the 50 most frequently reported words for each grade. A) grade 5 = optimal quality of movement, B) grade 4 = good quality of movement, C) grade 3 = mild movement dysfunction, D) grade 2 = moderate movement dysfunction, E) grade 1 = severe movement dysfunction.

would show more agreement.

The reason that the range of scores for some horses is so variable is unknown. For instance, horse 11 had the largest variation but from the comments, it is not clear why scores differed so much as many comments were similar. While horses 7 and 20, with the most consistent scores, presented the most behavioural and handling difficulties. Lower scores were generally accompanied by detailed comments of deficits in movement quality, while the comments for higher scores tend to be less specific such as "good trot".

The high levels of intra-rater reliability are similar to those achieved by veterinary physiotherapists in the FINFUN, a functional movement outcome measure for canines with neurological conditions [30], or in human physiotherapy for assessing balance [31] or gait [32]. Existing equine lameness scales also generally show moderate intra-rater reliability [19,33]. More experienced assessors and clear definitions of each grade for each movement appear to be associated with measures having higher intra-rater reliability [21,34,35]. The EQoMS's high intra-rater reliability means it could be effectively used by a single/solo clinician.

The inter-rater reliability for total scores was very poor, but when ranked there was moderate agreement as the assessors mostly put the horses in the same order. This finding is promising as it means they agreed on the relative quality of movement but not on the absolute amount of quality. The ability of equine clinicians to score lameness is also variable, with better agreement between observers on straight lines than circles and for forelimb lameness than hindlimb lameness [15]. The results of this study align with those of Starke and Oosterlinck [36] who

reported low inter-rater reliability for lameness scores between sound and mildly lame horses. As a concept quality of movement includes a broad range of features that even with completing the same online training course, there is subjectivity in what individuals' value. Diversity of opinions is not unique to the assessment of equine in-hand quality of movement, dressage, and other judged sports (diving, gymnastics, figure skating…). Delphi studies looking at equine welfare have found different people/ federations/ countries prioritise different aspects of welfare [37]. In contrast the FINFUN had very high inter-rater reliability [30] perhaps being condition specific and focussing on the level of independence or assistance helps to define the boundaries between grades enhancing inter-rater reliability.

The moderate correlations highlight that movement tests in different gaits (walk, trot, and canter), figures (straight line or circle) and surfaces (firm or soft) all have merit for assessing quality of movement, as they present different physical challenges that emphasise deficits in different aspects of movement. The importance of including canter in assessment has been previously established [17,38,39]. Similarly, different conclusions have been drawn from assessment of lameness at trot in a straight line only compared to a full lameness work up [40–42]. Performance at the trot in a straight line on a firm surface was not closely correlated with scores for canter on the soft surface, highlighting the need for assessment to be specific to the reported presenting problem/s. Gait made more difference than surface or direction, scores for walk on the firm and soft were more closely correlated than scores between walk and trot. Scores for the straight line and circle differed considerably, while scores for left and right rein for all circles at the same gait were closely related. This accords with previous evaluations of lameness and asymmetry on circles and straight lines [41,43,44]. Observing horses move in various contexts is particularly useful when visually assessing in the field [45]. While instrumented measures may be superior to the human eye at identifying small asymmetries [46], for the clinician in the field using progressive movement tests may help make a subtle issue visible; the level of asymmetry may increase from being below what can be visually detected by the human eye to becoming visually detectable [47,48]. Not forgetting, bilateral lameness, which may confound asymmetry measures [49]. Additionally, different movements, gaits and surfaces elicit different comments from the assessors, suggesting their inclusion allows the assessment of different features of movement quality, creating a more holistic assessment.

The comments accompanying the scores are a rich source of data which adds depth to the evaluation of the results. There was a notable absence of positive comments, despite the directives and scale using positive language. It may have been assumed that positive features should be present, and therefore, assessors neglected to mention them; this is supported by the much lower number of comments made when awarding a score of 4 or 5. Given the brain's negativity bias it is easier to be critical than to praise [50]. In any case, the language used to speak to owners is different to that used in clinical notes. It seems the assessors utilised the directives as many of the comments used words that are in the directives.

In line with previous studies [10,17] the equine physiotherapists in this study gave primary importance to lameness or asymmetry before attending to motor control aspects. This is reflected in lameness, limbs and the head being commented on most frequently. Starke and May [44] also showed eye gaze spent more time looking at the front of the horse, reflected in this study by commenting on the head more than the pelvis. The focus on the head may also tie in with the importance placed on signs of discomfort, of which the facial features are key indicators [51, 52]. Terms such as bend and stiffness may relate to aspects of movement beyond hip-hitching or head-nodding lameness and were commented on less frequently, suggesting these features may be viewed as secondary considerations. Lateral flexion varies with figure and gait but requires further investigation as a compensation strategy for specific movement dysfunctions [53–56]. Changes in spinal movement or stiffness may be related to pain [57,58].

Trot is the gait most frequently used to assess lameness, so it is not surprising that comments relating to lameness were more frequent for movement tests at trot than at walk or canter. Lameness was commented on more than twice as often when on a circle than on a straight line. The change in loading induced by the figure provokes asymmetry [49,54], which may increase the detection of subtle lameness. As expected, lateral bend was commented on for circles but not for the straight line, where the horse should ideally be straight/symmetrical.

It is recognised that different surfaces stress different tissues of the body more [59–61], which is reflected in the comments of the assessors. Non-contextualised comments about weakness, relating to motor control and muscle activity were more frequent for movement tests on soft surfaces. The soft surface may provide some cushioning to bony pathology with less comments about lameness on the soft surface [29].

While both professions considered lameness and asymmetry of primary importance the specialist equine veterinarians in this study used the words lame/ness and identified a limb three times more than the equine physiotherapists. On closer inspection the equine physiotherapists could be using alternative phrases like stepping short to describe lameness to avoid saying the horse is lame, because that is a more diagnostic phrase, and veterinary physiotherapists are not allowed to diagnose in horses [62]. So, they may be commenting on lameness, but modifying their choice of words. The use of different terminology has the potential for misinterpretation of findings between professions and horse owners.

While these six equine physiotherapists and specialist equine veterinarians did comment on different aspects of movement, they also shared many observations which shows the benefit of standardized movement tests and directives. Assessors from both professions commented on aspects of movement beyond basic lameness features, such as lateral bend, balance, co-ordination, rigidity, stiffness, tension, spinal mobility, discomfort, pain face and willingness. This supports the need for outcome measures with a broader scope than lameness identification. The equine physiotherapists tended to comment on a broader range of features, but within this small sample, they did not definitely focus more on motor control.

Like any subjective, ordinal outcome measure the EQoMS has numerous limitations. The EQoMS is intended to cover the key movements routinely assessed but it is challenging to balance comprehensiveness with conciseness. Originally, it was envisaged as a short test with the optional addition of patient-specific movements, including riding when relevant, to be interpreted as part of the whole examination, not in isolation. A survey of equine clinicians determined the movements to be included [3] and there is currently insufficient data to conclusively decide to exclude some movement tests. Modifications to assessing the transitions could be investigated further. A future study could look at removing the small circle and incorporating that into the lunge circle. A digital notes or app format may alleviate some of the concerns expressed around managing the length of the form. Furthermore, not all horses would complete the full battery; testing would be ceased if the horse was worsening or if at its current stage of rehabilitation may only perform certain movement tests.

Assessors found the transitions difficult to score; additional training may improve this. Most horses behaved well, but training or behavioural issues displayed by some horses negatively impacted the scoring process. Stepping back provoked the most negative behaviours; a few horses exhibited head shaking and one attempted to bite the handler. Some assessors noted that poor handling techniques interfered with their assessment. Using a consistent handler for all horses might improve reliability, as privately owned horses handled by their owner's introduced variability.

All horses in the study showed some degree of movement dysfunction in some movement tests and the sample may be skewed towards lower quality of movement. Until filming was complete it was unknown what range of quality would be presented. Selective recruitment of several performance horses with higher levels of soundness was

considered, but not feasible due to time constraints and project logistics. The study did not determine if the outcome measure had ceiling effects and clinicians may be conservative in declaring a horse free from issues. Arguably, optimal performance should be based on the individual horse's purpose, not an ideal standard. In this study, unlike in clinical practice, the assessors were not provided with detailed background history or performance expectations, which would influence perceptions of optimal movement.

The assessors performed their scoring by viewing video footage rather than seeing the horses live. Unfortunately, audio had to be removed from a very small number of video clips due to conversations between the camera operator and horse owner. There is no literature mentioning animal physiotherapy assessment via video, however physiotherapy successfully used telehealth during the COVID-19 pandemic and several human studies show comparable benefits to in-person physiotherapy [63]. Video assessment of lameness is considered less accurate than in-person assessment. Although intra and inter-rater agreement on lameness is higher when observing horses live [64] there are practical difficulties of having multiple assessors present in a live session. Use of high-quality video has advantages such as recruiting geographically diverse experts.

Further refinement is needed in the pursuit of optimal practice for assessing quality of movement in the field. Future research may involve testing with diverse groups of horses including performance and sport horses and clinical pilot testing to gather clinician feedback. The sample of horses used in this study may not be representative of the broader population for which the EQoMS is intended. Therefore, replication with horses undergoing rehabilitation or performance management, as well as those with milder or no movement dysfunction, is essential to evaluate the EQoMS's ability to differentiate quality of movement levels. Key features in video footage could be coded to refine the directives along with repeating assessments pre- and post-treatment to determine the tool's sensitivity to change. Since the current tests are time-consuming, they may be better suited for long-term monitoring with reassessment gaps of weeks rather than same-day repeats. A larger catalogue of data is needed to decide if some movement tests should be removed, or others added, and it would be beneficial to determine if individual tests can serve as standalone reassessments during treatment sessions. Following such refinements, inter-rater reliability could be retested with a larger sample of assessors, which may allow statistical analysis of individual movement tests as well as total scores.

Collaboration between equine clinicians and researchers is needed to continue refining the most appropriate features of the EQoMS for everyday use and greater efficiency. This grading system for subjective visual assessment is a starting point, however, more stratification is needed when visualising the junction between the idealised sound horse and those with movement dysfunction. Allowing scores for each individual movement test may reduce disagreement compared to trying to agree on one overall score. As discussion within and between professions is necessary to define the boundaries between grades perhaps focus groups or workshops could be conducted [65]. Presumably, similar processes have occurred for lameness scoring and dressage judging over the years, to move towards greater consensus on what features of movement should cause a score to be lower or higher.

Subjective measures are reliant on the proficiency of the clinician and are more susceptible to fluctuations than objective measures. Clinicians need supervision and training in using the scoring system, a thorough knowledge of administration guidelines, and clinical experience with the population under study. Future online training could include a testing element for calibrating both intra and inter-rater reliability, instead of self-evaluation of proficiency. Variation in scores and comments emphasises the subjective nature of movement quality and correspondingly how individuals value different aspects. In research assessing affective state of horses there tends to be more agreement on prominent indicators, than on subtle cues [66]. Despite all this the assessors agreed on the relative (ranking) quality of movement of horses in this study.

5. Conclusion

The EQoMS which combines numerical grading and descriptive comments developed for use in the field [10] was subject to reliability testing. Six specialist equine veterinarians and equine physiotherapists demonstrated excellent intra-rater reliability when using EQoMS. When scores were ranked, inter-rater reliability was moderate, however, there were some differences in comments between professions. The agreement on the relative quality of movement but disagreement on the absolute magnitude of the quality observed highlights the individual weighting placed on qualitative and quantitative elements. While further refinement is required to improve agreement on the boundaries between scores, the EQoMS provides a starting point for clinicians to grade their visual assessment of quality of movement.

Ethics

All authors agree that:

This research presents an accurate account of the work performed, all data presented are accurate and methodologies detailed enough to permit others to replicate the work.

This manuscript represents entirely original works and or if work and/or words of others have been used, that this has been appropriately cited or quoted and permission has been obtained where necessary.

This material has not been published in whole or in part elsewhere.

The manuscript is not currently being considered for publication in another journal.

That generative AI and AI-assisted technologies have not been utilized in the writing process or if used, disclosed in the manuscript the use of AI and AI-assisted technologies and a statement will appear in the published work.

That generative AI and AI-assisted technologies have not been used to create or alter images unless specifically used as part of the research design where such use must be described in a reproducible manner in the methods section.

All authors have been personally and actively involved in substantive work leading to the manuscript and will hold themselves jointly and individually responsible for its content.

Funding

This research was supported by an Australian Government Research and Training Program Scholarship Stipend.

CRediT authorship contribution statement

Conceptualisation, A.G.B., H.R., G.T. and R.L.; methodology, A.G.B., H.R., G.T. and R.L; validation, A.G.B.; formal analysis, A.G.B.; investigation, A.G.B. and M.D.; data curation, A.G.B.; writing—original draft preparation, A.G.B.; writing—review and editing, A.G.B., H.R., G.T., R.L and M.D; visualisation, A.G.B.; supervision, H.R., G.T. and R.L; project administration, H.R. All authors have read and agreed to the published version of the manuscript.

Declaration of Generative AI and AI-assisted technologies in the writing process

During the preparation of this work the primary author used ChatGPT (version 3.5) to summarise project notes before writing the first draft. After using this tool, the authors reviewed and edited the content as needed and take full responsibility for the content of the publication.

Declaration of Competing Interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: Annette Bowen reports financial support was provided by Australian Government Research and Training Program Scholarship Stipend. If there are other authors, they declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

Thank you to the Charles Sturt University horses and equine staff, the privately owned horses and their humans, Courtney Malsem, Marina Douglas, and Bachelor of Equine Science students for assistance during data collection in Wagga Wagga. Thank you to Gail Fuller and Deanna Duffy, the Survey Monkey® administrators for Charles Sturt University's (CSU) Spatial Data Analysis Network (SPAN). Those acknowledged below provided their permission to be mentioned as expert participants: Dr A. Kaneps (Kaneps Equine Sports Medicine and Surgery, USA), Dr K. Neil (Sporthorse Veterinary Specialist, Australia), S. Palmer (The Horse Physio UK), J. Paul (RAMP, Scotland), Dr C. Steel (Hong Kong Jockey Club, HK), and M. Teeling (Equine-librium College, South Africa).

Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at doi:10.1016/j.eqre.2025.100023.

References

[1] S. Dyson, Can lameness be graded reliably? Equine Vet. J. 43 (2011) 379–382.
[2] L. Goff, Physiotherapy assessment for the equine athlete, Vet. Clin. North Am.: Equine Pract. 32 (2016) 31–47.
[3] A.G. Bowen, G. Tabor, R. Labens, H. Randle, Visually assessing equine quality of movement: a survey to identify key movements and patient-specific measures, Animals 13 (2023) 2822.
[4] W. Brookes, R. Payne, R. Lea, Reporting outcome measures in veterinary physiotherapy with particular reference to the treatment of canine and equine joint cases in the UK, Vet. Rec. 194 (2024).
[5] S. Egan, P. Brama, D. McGrath, Irish equine industry stakeholder perspectives of objective technology for biomechanical analyses in the field, Animals 9 (2019) 539.
[6] H. Randle, M. Steenbergen, K. Roberts, A. Hemmings, The use of the technology in equitation science: a panacea or abductive science? Appl. Anim. Behav. Sci. 190 (2017) 57–73.
[7] J. Thirkell, R. Hyland, A survey examining attitudes towards equine complementary therapies for the treatment of musculoskeletal injuries, J. Equine Vet. Sci. 59 (2017) 82–87.
[8] S. Brownlee, K. Chalkidou, J. Doust, A.G. Elshaug, P. Glasziou, I. Heath, et al., Evidence for overuse of medical services around the world, Lancet 390 (2017) 156–168.
[9] A.C. Traeger, R.N. Moynihan, C.G. Maher, Wise choices: making physiotherapy care more valuable, J. Physiother. 63 (2017) 63–65.
[10] Bowen A., Tabor G., Labens R., Randle H. Development of a grading system for visually assessing equine quality of movement. Journal of Equine Rehabilitation. 2024;submitted.
[11] de Vet H.C.W. Measurement in medicine: a practical guide. Cambridge: Cambridge University Press; 2011.
[12] F.M. Serra Bragança, M. Rhodin, P.R. van Weeren, On the brink of daily clinical application of objective gait analysis: what evidence do we have so far from studies using an induced lameness model? Vet. J. 234 (2018) 11–23.
[13] C.M. Crecan, C.P. Peştean, Inertial sensor technologies—their role in equine gait analysis, a review, Sensors 23 (2023) 6301.
[14] S.K. Reed, J. Kramer, L. Thombs, J.B. Pitts, D.A. Wilson, K.G. Keegan, Comparison of results for body-mounted inertial sensor assessment with final lameness determination in 1,224 equids, J. Am. Vet. Med. Assoc. 256 (2020) 590–599.
[15] M. Hammarberg, A. Egenvall, T. Pfau, M. Rhodin, Rater agreement of visual lameness assessment in horses during lungeing, Equine Vet. J. 48 (2016) 78–82.
[16] K.G. Keegan, E.V. Dent, D.A. Wilson, J. Janicek, J. Kramer, A. Lacarrubba, et al., Repeatability of subjective evaluation of lameness in horses, Equine Vet. J. 42 (2010) 92–97.
[17] A. Dubaniewicz-Pearce, G. Tabor, E. Davies, An exploration of visual gait assessment of horses by physiotherapists, J. Equine Rehabil. (2025) 100020.
[18] V. Venek, S. Kranzinger, H. Schwameder, T. Stöggl, Human movement quality assessment using sensor technologies in recreational and professional sports: a scoping review, Sensors 22 (2022) 4786.
[19] C.J. Fuller, B.M. Bladon, A.J. Driver, A.R.S. Barr, The intra- and inter-assessor reliability of measurement of functional outcome by lameness scoring in horses, Vet. J. 171 (2006) 281–286.
[20] Serra Bragança, F.M. Brommer, H. van den Belt, A.J.M. Maree, J.T.M. van Weeren, P.R. van Oldruitenborgh-Oosterbaan, MMS. Subjective and objective evaluations of horses for fit-to-compete or unfit-to-compete judgement, Vet. J. 257 (2020) 105454.
[21] T.J.P. Spoormakers, E.A.M. Graat, F.M. Serra Bragança, P.R.V. Weeren, H. Brommer, Rater agreement for assessment of equine back mobility at walk and trot compared to quantitative gait analysis, PLOS One 16 (2021) e0252536.
[22] A. Meier, M. de Laat, C. Pollitt, D. Walsh, J. McGree, D. Reiche, et al., A "modified Obel" method for the severity scoring of (endocrinopathic) equine laminitis, PeerJ 7 (2019) e7084.
[23] M. Marcantonio Coneglian, T. Duarte Borges, S.H. Weber, H. Godoi Bertagnon, P. V. Michelotto, Use of the horse grimace scale to identify and quantify pain due to dental disorders in horses, Appl. Anim. Behav. Sci. 225 (2020) 104970.
[24] S. Dyson, K. Thomson, L. Quiney, A. Bondi, A.D. Ellis, Can veterinarians reliably apply a whole horse ridden ethogram to differentiate nonlame and lame horses based on live horse assessment of behaviour? Equine Vet. Educ. 32 (2020) 112–120.
[25] T. Pfau, K. Noordwijk, M.F. Sepulveda Caviedes, E. Persson-Sjodin, A. Barstow, B. Forbes, et al., Head, withers and pelvic movement asymmetry and their relative timing in trot in racing Thoroughbreds in training, Equine Vet. J. 50 (2018) 117–124.
[26] I. Wolframm, Let them be the judge of that: bias cascade in elite dressage judging, Animals 13 (2023) 2797.
[27] T. Whitaker, H. Randle, A. Mills, Start time and effect of order in the dressage phase of affiliated British Eventing (BE100 Level) competition, J. Vet. Behav. 8 (2013) e23–e24.
[28] B. Forbes, W. Ho, R.S.V. Parkes, M.F. Sepulveda Caviedes, T. Pfau, D.R. Martel, Associations between racing Thoroughbred movement asymmetries and racing and training direction, Animals 14 (2024) 1086.
[29] Clayton H.M. The dynamic horse; A biomechanical guide to equine movement and performance. Mason, MI: Sport Horse Publications; 2004.
[30] A.F. Boström, H.K. Hyytiäinen, P. Koho, S. Cizinauskas, A.K. Hielm-Björkman, Development of the Finnish neurological function testing battery for dogs and its intra- and inter-rater reliability, Acta Vet. Scand. 60 (2018) 56.
[31] S. Downs, J. Marquez, P. Chiarelli, The Berg Balance Scale has high intra- and inter-rater reliability but absolute reliability varies across the scale: a systematic review, J. Physiother. 59 (2013) 93–99.
[32] G. Chamorro-Moriana, C. Ridao-Fernández, J. Ojeda, M. Benítez-Lugo, J. L. Sevillano, Reliability and validity study of the Chamorro assisted gait scale for people with sprained ankles, walking with forearm crutches, PLOS One 11 (2016) e0155225 (e).
[33] M. Hewetson, R.M. Christley, I.D. Hunt, L.C. Voute, Investigations of the reliability of observational gait analysis for the assessment of lameness in horses, Vet. Rec. 158 (2006) 852–858.
[34] J.J. Brunnekreef, C.J.T. van Uden, S. van Moorsel, J.G.M. Kooloos, Reliability of videotaped observational gait analysis in patients with orthopedic impairments, BMC Musculoskelet. Disord. 6 (2005) 17.
[35] R.W. Moran, A.G. Schneiders, K.M. Major, S.J. Sullivan, How reliable are Functional Movement Screening scores? A systematic review of rater reliability, Br. J. Sports Med. 50 (2016) 527–536.
[36] S.D. Starke, M. Oosterlinck, Reliability of equine visual lameness classification as a function of expertise, lameness severity and rater confidence, Vet. Rec. 184 (2019) 8.
[37] J.M. Williams, L.C. Berg, H.M. Clayton, K. Kirsch, D. Marlin, H. Randle, et al., A Delphi study to determine international and national equestrian expert opinions on domains and sub-domains essential to managing sporthorse health and welfare in the Olympic disciplines, Animals 13 (2023) 3404.
[38] S. Dyson, R. Murray, Pain associated with the sacroiliac joint region: a clinical study of 74 horses, Equine Vet. J. 35 (2003) 240–245.
[39] S. Dyson, Equine performance and equitation science: clinical issues, Appl. Anim. Behav. Sci. 190 (2017) 5–17.
[40] S.K. Reed, J. Kramer, L. Thombs, J.B. Pitts, D.A. Wilson, K.G. Keegan, Comparison of results for body-mounted inertial sensor assessment with final lameness determination in 1,224 equids, J. Am. Vet. Med. Assoc. 256 (2020) 590–599.
[41] M. Rhodin, L. Roepstorff, A. French, K.G. Keegan, T. Pfau, A. Egenvall, Head and pelvic movement asymmetry during lungeing in horses with symmetrical movement on the straight, Equine Vet. J. 48 (2016) 315–320.
[42] S. Dyson, L. Greve, Subjective gait assessment of 57 sports horses in normal work: a comparison of the response to flexion tests, movement in hand, on the lunge, and ridden, J. Equine Vet. Sci. 38 (2016) 1–7.
[43] A. Byström, A. Hardeman, F. Braganca, L. Roepstorff, J.-H. Swagemakers, P. van Weeren, et al., Differences in equine spinal kinematics between straight line and circle in trot, Sci. Rep. 11 (2021).
[44] S.D. Starke, S.A. May, Expert visual assessment strategies for equine lameness examinations in a straight line and circle: a mixed methods study using eye tracking, Vet. Rec. 191 (2022).
[45] L. Greve, S. Dyson, What can we learn from visual and objective assessment of non-lame and lame horses in straight lines, on the lunge and ridden? Equine Vet. Educ. 32 (2020) 479–491.
[46] T. Pfau, P. Reilly, How low can we go? Influence of sample rate on equine pelvic displacement calculated from inertial sensor data, Equine Vet. J. (2020).
[47] S. Dyson, Recognition of lameness: man versus machine, Vet. J. 201 (2014) 245–248.

[48] R.S.V. Parkes, R. Weller, A.M. Groth, S. May, T. Pfau, Evidence of the development of 'domain-restricted' expertise in the recognition of asymmetric motion characteristics of hindlimb lameness in the horse, Equine Vet. J. 41 (2009) 112–117.
[49] L. Greve, T. Pfau, S.J. Dyson, Alterations in body lean angle in lame horses before and after diagnostic analgesia in straight lines in hand and on the lunge, Vet. J. 239 (2018) 1–6.
[50] T.A. Ito, J.T. Larsen, N.K. Smith, J.T. Cacioppo, Negative information weighs more heavily on the brain: the negativity bias in evaluative categorizations, J. Personal. Soc. Psychol. 75 (1998) 887–900.
[51] K.B. Gleerup, C. Lindegaard, Recognition and quantification of pain in horses: a tutorial review, Equine Vet. Educ. 28 (2016) 47–57.
[52] S. Dyson, J.M. Berger, A.D. Ellis, J. Mullard, Can the presence of musculoskeletal pain be determined from the facial expressions of ridden horses (FEReq)? J. Vet. Behav. 19 (2017) 78–89.
[53] A. Egenvall, H. Engström, A. Byström, Back motion in unridden horses in walk, trot and canter on a circle, Vet. Res. Commun. (2023).
[54] L. Greve, T. Pfau, S.J. Dyson, Thoracolumbar movement in sound horses trotting in straight lines in hand and on the lunge and the relationship with hind limb symmetry or asymmetry, Vet. J. 220 (2017) 95–104.
[55] A.M. Hardeman, A. Byström, L. Roepstorff, J.H. Swagemakers, P.R. van Weeren, F. M. Serra Bragança, Range of motion and between-measurement variation of spinal kinematics in sound horses at trot on the straight line and on the lunge, PLOS One 15 (2020) e0222822 (e).
[56] E. Zetterberg, E. Persson-Sjodin, J. Lundblad, E. Hernlund, M. Rhodin, Prevalence of movement asymmetries in high-performing riding horses perceived as free from lameness and riders' perception of horse sidedness, PLOS One 19 (2024) e0308061.
[57] L. Greve, S. Dyson, T. Pfau, Alterations in thoracolumbosacral movement when pain causing lameness has been improved by diagnostic analgesia, Vet. J. 224 (2017) 55–63.
[58] K.R. Holm, J. Wennerstrand, U. Lagerquist, P. Eksell, C. Johnston, Effect of local analgesia on movement of the equine back, Equine Vet. J. 38 (2006) 65–69.
[59] A. Barstow, J. Bailey, J. Campbell, C. Harris, R. Weller, T. Pfau, Does 'hacking' surface type affect equine forelimb foot placement, movement symmetry or hoof impact deceleration during ridden walk and trot exercise? Equine Vet. J. 51 (2019) 108–114.
[60] J.L. Mendez-Angulo, A.M. Firshman, D.M. Groschen, P.J. Kieffer, T.N. Trumble, Impact of walking surface on the range of motion of equine distal limb joints for rehabilitation purposes, Vet. J. 199 (2014) 413–418.
[61] T. Ursini, K. Shaw, D. Levine, H. Steve Adair, J. Richards, Electromyography of the multifidus muscle in horses trotting over firm and soft surfaces, J. Equine Rehabil. 1 (2023) 100004.
[62] C.M. McGowan, S. Cottriall, Introduction to equine physical therapy and rehabilitation, Vet. Clin. North Am.: Equine Pract. 32 (2016) 1–12.
[63] S. Mani, S. Sharma, B. Omar, A. Paungmali, L. Joseph, Validity and reliability of Internet-based physiotherapy assessment for musculoskeletal disorders: a systematic review, J. Telemed. Telecare 23 (2017) 379–391.
[64] Ross M.W. Movement. In: Ross MW, Dyson SJ, editors. Diagnosis and management of lameness in the horse. 2nd ed. St Louis, Missouri: Elsevier Saunders; 2011;64-80.
[65] Haussler K.K. Is it time for a paradigm shift in veterinary healthcare? Veterinary Compedium, 2024.
[66] T. Bornmann, H. Randle, J. Williams, Investigating equestrians' perceptions of horse happiness: an exploratory study, J. Equine Vet. Sci. 104 (2021) 103697.