Journal of Veterinary Behavior 9 (2014) 22–29

Contents lists available at ScienceDirect

Journal of Veterinary Behavior

journal homepage: www.journalvetbehavior.com

Research

Assessing ridden horse behavior: Professional judgment and physiological measures

Carol Hall*, Rachel Kay, Kelly Yarnell

School of Animal, Rural and Environmental Sciences, Nottingham Trent University, Brackenhurst Campus, Southwell, Nottinghamshire, United Kingdom

A R T I C L E I N F O

Article history:
Received 12 March 2013
Received in revised form
19 August 2013
Accepted 12 September 2013
Available online 2 December 2013

Keywords:
horse
stress
ridden behavior
cortisol
eye temperature

A B S T R A C T

The assessment of ridden horse behavior by 12 equestrian professionals (riding instructors n = 4, riders n = 4, veterinarians n = 4) was compared with observed behavior and physiological measures (salivary cortisol and eye temperature). Horses (n = 10) were ridden at walk, trot, and canter in a predefined test of approximately 2-3 minutes. Video footage of the ridden test (RT) was analyzed using Observer XT 10 and duration of behavioral states/events recorded. Saliva was collected in the stable, after the warm-up (WU) and at 0, 5, 15, 30, and 60 minutes after the RT. The saliva was analyzed for cortisol (enzyme-linked immunosorbent assay) and the difference between minimum and maximum concentration (ng/mL) and associated sample times recorded. Eye temperature was measured using an infrared thermal camera (MobIR M8), static images (stable, after WU, after RT), and video footage (WU and RT) with maximum eye temperatures derived from set intervals. Mean maximum eye temperatures during ridden work were calculated. Video footage of the RT was observed by the 12 equestrian professionals who each scored the horses on 7 performance parameters derived from the Fédération Equestre Internationale rules for dressage events and the training scale of the German National Equestrian Federation (relaxation, energy, compliance, suppleness, confidence, motivation, and happiness). These scores were compared with behavioral and physiological measures and correlations investigated (Spearman's rank order correlation). Higher percentage durations of high head carriage (ranging from 0 to 50.75% of RT) and the nose carried at an angle in front of the vertical (0%-74.29% of RT) correlated with overall less favorable assessment by the equestrian professionals ($P < 0.05$) and only the instructors associated neutral head carriage (32.76%-91.92% of RT) and vertical nasal angle (0.97%-68.90% of RT) as a positive sign ($P = 0.03$ and $P = 0.04$, respectively). Increases in salivary cortisol positively correlated with the duration of low head carriage ($P < 0.05$), suggesting that this way of going increased the demands placed on the horse. Increased eye temperature positively correlated with duration of nose carried behind the vertical when ridden ($P = 0.02$) and negatively correlated with duration of nose carried in front of the vertical ($P = 0.01$). Some discrepancy between physiological evidence and professional assessment of ridden horse behavior was evident as were differences between groups of professionals. Further evaluation of the association between behavioral signs and physiological measures is now required to ensure that the assessment of ridden horse performance is based on valid and consistent measures.

Introduction

Ridden horse assessment is generally based on subjective judgment of observed behavior. In addition to physical soundness, veterinarians and other equestrian professionals need to be able to identify signs of mental distress in ridden horses (Ödberg, 1987), which can contribute to poor performance, health, and behavioral problems. Within international equestrian sport, there is much debate surrounding the stress associated with training methods, in particular in relation to head and neck position (van Breda, 2006; von Borstel et al., 2009; McGreevy et al., 2010). In response, the Fédération Equestre Internationale (FEI, 2009) published guidelines advocating harmonious education of the horse resulting in overt behavior indicative of "submission" and the development of the horse into a "happy athlete." The guidelines state that the nasal plane should at all times be positioned in front of the vertical (FEI, 2009). However, there is evidence that overflexion of the neck (and

R.K. and K.Y. contributed equally to the work.

* Address for reprint requests and correspondence: Carol Hall, PhD, School of Animal, Rural and Environmental Sciences, Nottingham Trent University, Brackenhurst Campus, Southwell, Nottinghamshire, NG25 0QF, UK. Tel.: +44 (0) 1158 485212.
E-mail address: carol.hall@ntu.ac.uk (C. Hall).

1558-7878/$ – see front matter
http://dx.doi.org/10.1016/j.jveb.2013.09.005

the consequent positioning of the nasal plane behind the vertical) is still considered by some as a positive sign in the ridden horse (McGreevy et al., 2010). For example, during official stallion performance tests, higher scores for ride ability were awarded to horses that were ridden with their nose line predominantly behind the vertical than to those ridden with their nose line at the vertical (König von Borstel et al., 2011). Behavioral events such as grinding the teeth and agitation of the tail are considered signs of nervousness, tension, or resistance on the part of the horse (FEI, 2009). In their judgment of ridden horse behavior, equestrian professionals will rely on overt behavioral signs, but the relative importance attributed to specific signs is unlikely to be consistent across professional disciplines. To determine whether the interpretation of behavioral signs used in such judgments is justifiable, more objective evidence is required.

In ridden work, physiological measures of mental stress are confounded by the effect of the physical demands of exercise (Marlin and Nankervis, 2002) as well as problems in distinguishing between excitement/arousal and fear/anxiety. Factors such as fitness and age need to be taken into consideration if hormonal stress responses are to be used as a means of interpreting behavioral signs. However, in low-level exercise scenarios, findings suggest that anxiety-provoking situations produce increases in salivary cortisol concentration (Schmidt et al., 2010). Training per se affects baseline plasma cortisol concentration in riding horses (Fazio et al., 2006) and subsequent cortisol response to ridden events is affected by previous training (Fazio et al., 2008). With due consideration to such factors, salivary cortisol response to low levels of established ridden work may provide evidence of the demand characteristics of some aspects of ridden horse behavior.

Circulatory changes are a more immediate physiological response to sympathetic activation (preparing the animal for flight or fight) and are associated with changes in surface temperature. These can be measured using infrared thermography and have been used as a noninvasive means of assessing stress responses in animals (Stewart et al., 2005). In particular, eye temperature has been shown to increase in some species in response to potentially distressing procedures (Cook et al., 2006; Stewart et al., 2007). In the horse, increases in eye temperature were found to correlate with increases in salivary cortisol during a potentially aversive procedure (clipping) (Yarnell et al., 2013) and higher eye temperatures were recorded when horses were lunged in a training aid (Pessoa) than when lunged without (Hall et al., 2011). Eye temperature offers a means of confirming (or not) the interpretation of behavioral signs in the ridden horse.

The aim of this study was to compare the judgment of ridden horse behavior made by equestrian professionals with evidence from both observed behavior and physiological measures. Associations between scoring and the duration of observed behavior and between behavior and physiological responses were investigated. The long-term aim is to identify behavioral signs in the ridden horse that are indicative of mental state.

Materials and methods

Horses

The study involved 10 riding horses (6 geldings, 4 mares) and 10 experienced riders (who rode the horse regularly). Horse heights ranged from 158 to 173 cm, their ages from 5-20 years (mean age 12.9 years), and they were thoroughbred and thoroughbred × warmblood cross in type. The current roles of the horses were either predominantly competitive (n = 6), leisure (n = 2), or riding school work (n = 2). All horses were in work at the time of the study (ridden 5 days/week) and were at levels of fitness befitting their current workload. This ranged from fit for regular hacking (walk, trot, and canter) to competition fit (British Eventing novice level). The trial was carried out in the United Kingdom during July and the horses were field kept (n = 4), or on a combined stable/field turnout regime, stabled during the day (n = 2) or stabled overnight (n = 4).

Experimental design

Each horse/rider combination completed a ridden trial consisting of a warm-up (WU) and ridden test (RT). Behavior was recorded from the video footage of the RT; this footage was also assessed by equestrian professionals (n = 12). Physiological measures were taken throughout the trial (salivary cortisol and eye temperature). Handling behavior was scored at sampling sessions before and after the RT to assess the impact of the ridden work on general behavior. See Table 1 for details of the sampling times for each measure. The study was conducted in accordance with the Nottingham Trent University's ethical review process.

Ridden trials

The ridden trials were all conducted at the horses' home yards and ridden work took place outside between 10.30 and 15.00 hours. The horse was prepared for ridden work (normal tack with no martingales, draw reins, or other schooling aids; 4 horses wore flash nosebands, the rest wore plain cavesson nosebands), led to the riding area, mounted, and warmed up (WU) for 10-20 minutes (including work in walk, trot, and canter on both reins). Each horse was ridden in a fenced arena on an all-weather surface in an area of

Table 1
Sampling times for each measure taken during the study

Measure	Before ridden work	WU	After WU/before RT	RT	After ridden work
Ridden behavior				% duration of behavior recorded from video	
Saliva sampling (cortisol)	Sampled before preparation for ridden work		Sampled after WU		Sampled 0, 5, 15, 30, and 60 minutes after RT
Eye temperature (static IRT)	Maximum eye temperature recorded		Maximum eye temperature recorded		Maximum eye temperature recorded (0, 5, 15, 30, and 60 minutes after RT)
Eye temperature (video IRT)		Maximum eye temperature recorded at 3 set points during WU		Maximum eye temperature recorded at 7 set points during RT	
Scoring by professionals				RT videos scored on 7 parameters	
Handling test	Behavior scored from video				Behavior scored from video (0 and 60 minutes after RT)

IRT, infrared thermography; RT, ridden test; WU, warm-up.

approximately 20 × 40 m. No other horses were present during the trials. The RT consisted of set movements as described in Table 2.

The RT was recorded with a Canon Legria FS306 digital video camera (hand held) positioned at 1 of the short sides of the arena. Recording began when the horse passed approximately halfway down the long side of the arena approaching the video camera and ceased when the horse came to a halt at the end of the test. The person filming the RT directed the camera to follow the horse around to keep it centered in the video.

Behavior during ridden test

The video footage (RT) was analyzed using Observer XT 10 behavioral analysis software. See Table 3 for a list of behavioral categories, associated states, and descriptors (adapted from Weeks, 1996; Kaiser et al., 2006; Heleski et al., 2009; Hall et al., 2012). The duration of each behavioral state for each separate category was calculated using Observer XT 10 as a percentage of the duration of the RT for each horse. When specific behaviors were not visible, this was recorded to evaluate relative visibility.

Physiological measures

Saliva collection and cortisol analysis

Saliva was collected using Salivettes (Sarstedt, UK) designed for human use and modified by the addition of a cotton thread to enable the human handler to keep hold of the Salivette while the horse chewed. The swab was placed in the horse's oral cavity at the height of the third premolar in the maxilla for approximately 30 seconds. Salivettes were refrigerated at 4°C for no longer than 2 hours before they were transferred and frozen at −20°C until analysis. The saliva was analyzed for cortisol using a commercially available enzyme-linked immunosorbent assay (DRG Diagnostics). A total of 2 salivary cortisol assays were carried out for this study. The % coefficient of variation of means for high control was 3.93% and the coefficient of variation of means for low control was 7.36%. The mean interassay coefficient of variation was 5.65%. Cortisol concentration was calculated for each horse/sample time (ng/mL).

Table 2
Ridden test

Pace	Movements
Trot	
1	Start halfway down the long side facing the person filming and begin by trotting down the long side toward them (working trot)
2	Carry on trotting through the next short side[IRT]
3	Change the rein through the long diagonal, with a transition to and from walk in the centre of the arena (you should walk for approximately 1 horse's length)[IRT]
4	Continue around the short side of the arena
Canter	
5	Pick up canter anywhere on the short side (or in the corner)[IRT]
6	Proceed in working canter down the next long side and through the following short side
7	Change the rein through the long diagonal, with change of lead through trot in the center; you should trot for 2 to 4 strides[IRT]
8	Continue in working canter down the next long side until halfway, then ride downward transitions to trot[IRT] and walk
Walk	
9	Continue in medium walk down the long side and through the next short side
10	Change the rein through the long diagonal, with a free walk on a long rein[IRT]
11	Return to medium walk at the end of the diagonal and continue through the short side
12	Ride a downward transition to halt halfway through the short side [IRT]

[IRT], the points at which maximum eye temperature was recorded.

Table 3
Behavior recorded during ridden test with descriptors

Behavioral category	Variation of behavior within category	Description of behavioral variations
Head position/ nasal plane	Vertical	Nasal plane approximately vertical
	Behind vertical	Nasal plane at least 10° behind the vertical
	In front of vertical	Nasal plane at least 10° in front of the vertical
Head/neck carriage	High	Nose above the withers
	Neutral	Nose between withers and abdominal line
	Low	Nose below abdominal line
Head movement	Still central	Absence of lateral cervical flexion, head in line with rest of horse's body
	Still turned to left	Lateral cervical flexion to the left, not in line with rest of horse's body
	Still turned to right	Lateral cervical flexion to the right, not in line with rest of horse's body
	Still tilted left	Nasal midline not perpendicular to ground; deviation of angle of nasal midline to left
	Still tilted right	Nasal midline not perpendicular to ground; deviation of angle of nasal midline to right
	Tossing vertically	The horse moves the head in a quick forward–upward motion
	Tossing laterally	The horse moves the head in a quick lateral motion
Ear position	Both forward	Both ears immobile with pinna pointed forward
	Both back relaxed	Both ears turned backward without being flattened
	Both back (pinned)	Both ears are flattened backwards
	Right front, left rear	Right ear points forward, left ear turned backward
	Left front, right rear	Left ear points forward, right ear turned backward
Mouth	Still and shut	
	Still and open (tongue not visible)	Mouth open and visible space between upper and lower jaws (tongue not visible)
	Still and open (tongue visible)	Mouth open and visible space between upper and lower jaws (tongue visible)
	Moving	
Tail	Swinging with body movement	
	Swishing purposefully	Strong lateral and dorsoventral movements of the tail beyond that of simple rhythmic swinging
Salivation	Saliva not visible	
	Saliva visible	
Auditory signals	Snorting	Sound produced upon forceful quick exhalation of less than 1-second duration
	Vocalization	Loud, prolonged call, typically 1 to 3 seconds, beginning at high pitch and ending at lower pitch

Eye temperature

Eye temperature throughout the study was recorded using a MobIR M8 thermal camera (Wuhan Guide Infrared Company Ltd., China). Ambient temperature was recorded at each sampling time. Static thermal images from both left and right eye were captured at a distance of 1 meter ± 50 cm from the horse at a 90° angle immediately before saliva sampling. Images were uploaded to analytical software (Guide IrAnalyser) and the maximum temperature for the eye (from within the palpebral fissure around the entire eyelid margin from the lateral commissure to the lacrimal caruncle) recorded. No significant difference between right and left eye temperatures was found and mean maximum eye temperature (°C) was calculated for each horse/sample time (stable, post WU, post RT) for subsequent analyses.

During WU and RT video footage was collected using the same MobIR M8 thermal camera with an added 30-mm Tele lens, hand held adjacent to the position of the video camera. The footage was uploaded to the Guide IrAnalyser software and eye temperature (from the visible eye only) was extracted as for the static images. Maximum eye temperature was recorded at 3 points within WU (start; after 1 minute; after 5 minutes). Maximum eye temperature was recorded in RT at 7 set points (identified by^{IRT} in Table 2). Mean values were calculated for WU and RT.

Professional judgment

The video footage of the RT was assessed by 12 professionals, all of whom had ridden/trained horses for at least 10 years and whose primary role was described as either veterinarian (n = 4), riding instructor (n = 4), or rider (n = 4). These professionals were recruited from staff associated with the university and from nearby veterinary and academic institutions. The order in which the horses were viewed was different for each participant. The clips were played back on a PC/laptop and after watching each clip, the horse was assessed on 7 parameters, on a scale of 1 to 5. These parameters (Table 4) were derived from the FEI Rules for Dressage Events (FEI, 2009) and the training scale of the German National Equestrian Federation (1997).

Mean scores were calculated for each horse (for each parameter and overall) from all scorers together and also separately from veterinarians, instructors, and riders.

Handling test

The behavior of each horse was recorded at each saliva collection sampling time using a Canon Legria FS306 digital video camera. The camera was hand held and the operator was located at a distance of approximately 4 m from the horse. Video footage was uploaded onto a laptop computer and viewed using Windows media player. Video clips from three sampling times (the initial sampling time pre-RT, immediately after the RT, and 60 minutes after the RT) were used to score signs of relaxation/anxiety in each horse and the potential effect of the RT on general anxiety level. Each video clip was numbered (1-30) to allow identification of horse and stage of trial but then arranged in a random order for scoring purposes. The scoring scale used and a description of the behavior associated with each score is shown in Table 5.

Table 4
The parameters used for subjective evaluation of the horse in the ridden test and related scales (adapted from the scales of training of the German National Equestrian Federation, 1997 and the FEI Rules for Dressage Events, 2009)

Parameter	Score 1	Score 2	Score 3	Score 4	Score 5
Relaxation	Very relaxed	Quite relaxed	Neither relaxed nor tense	Quite tense	Very tense
Energy	Very energetic	Quite energetic	Neither energetic nor lethargic	Quite lethargic	Very lethargic
Compliance	Very compliant	Quite compliant	Neither compliant nor resistant	Quite resistant	Very resistant
Suppleness	Very supple	Quite supple	Neither supple nor stiff	Quite stiff	Very stiff
Confidence	Very confident	Quite confident	Neither confident nor timid	Quite timid	Very timid
Motivation	Very motivated	Quite motivated	Neither motivated nor unmotivated	Quite unmotivated	Very unmotivated
Happiness	Very happy	Quite happy	Neither happy nor unhappy	Quite unhappy	Very unhappy

Table 5
Scale used to assess level of relaxation (activity level) relating to response to procedure and behaviors associated with each score

Score	Type of response	Associated behaviors
1	Very relaxed	No movement apart from mouth; minimal ear movement. No attempt to avoid procedure.
2	Quite relaxed	Slow movements of head/ears. No attempt to avoid procedure.
3	Alert/interested	Ears forward. Head raised. Looking toward handler/procedure. Interest but no attempt to avoid.
4	Anxious	Head and neck moving away. Ears back/moving rapidly. No body movement. Initial attempt to avoid procedure.
5	Frightened/avoidance	Movement of head, neck, and body. Abrupt movement away. Repeated attempts to avoid procedure.

The 30 video clips were viewed and scored by 20 undergraduate students on an equine degree course and their tutor. Inter-observer variability was calculated (Cronbach's Alpha) to assess the reliability of the scoring. Mean behavior scores were then calculated for each horse/sample time combination.

Data analyses

All statistical analyses were carried out using IBM SPSS 19.

Non-parametric analyses were used throughout as the majority of data sets varied significantly from the normal distribution (Kolmogorov-Smirnov test: $P < 0.05$).

The overall effect of sample time on salivary cortisol concentration, eye temperature, and behavior during handling was assessed (Friedman test; Wilcoxon signed rank test). The effect of sex, housing, noseband type, and current role on these measures was tested (Mann-Whitney U test; Kruskal-Wallis test), as was their correlation with the age of the horse (Spearman rank order correlation).

To assess the impact of the ridden work on cortisol concentration, sample times associated with maximum and minimum concentration values were identified for each horse and the difference between them calculated for each horse for comparison with professional judgment, eye temperature, and behavioral data. Mean maximum eye temperatures were calculated for both static images (stable, post-WU, and post-RT) and video (WU and RT). Correlations between cortisol, eye temperature and the percentage durations of observed ridden behavior were assessed (Spearman rank order correlation).

Mean performance scores awarded for the RT by the professionals were calculated for each parameter (see Table 4) and overall. Mean scores awarded by each group of professionals were calculated separately to enable the comparison of the assessment by the different groups (Spearman rank order correlation). Correlation between these scores, recorded behavior, salivary cortisol, and eye temperature was assessed (Spearman rank order correlation).

In the handling test, the effect of sample time on mean behavior scores was assessed (Friedman test) as was the correlation between mean maximum eye temperature and mean behavior score (Spearman rank order correlation).

Results

Salivary cortisol

The cortisol response of individual horses varied and there was no consistent effect of sample time on cortisol concentration. However, some horses (n = 5) had minimum cortisol concentration

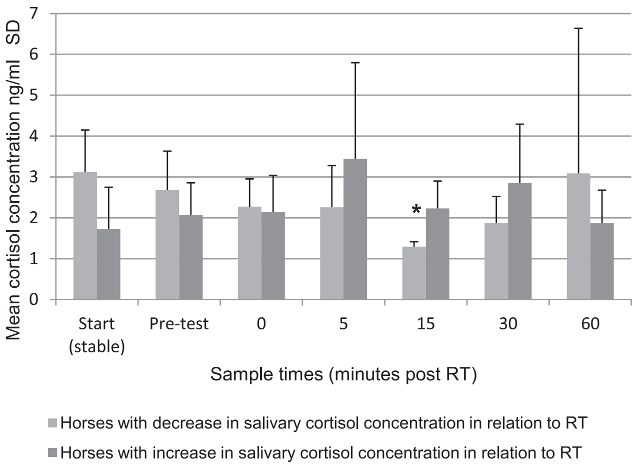

Figure. Mean cortisol concentration at each sample time for the 5 horses (1, 2, 5, 7, 8) that had a relative decrease in cortisol concentration with minimum concentration at 15 minutes after the ridden test (RT) (significantly lower than at any other sample time $P < 0.05$ as signified by *) and for the 5 horses (3, 4, 6, 9, 10) that had varying levels of increase in cortisol concentration after the RT (but no significant effect of sample time was found for this group) (± SD).

at 15 minutes after RT and maximum at either the initial sample time (stable: n = 3), 60 minutes after RT (n = 1), or immediately before the RT (n = 1). These horses showed a reduction in cortisol concentration associated with the ridden work, whereas others (n = 5) had an increase in cortisol concentration in response to ridden work (of varying amounts and at slightly differing times). In the horses showing a decrease, 1 horse (horse 7) was recorded as having maximum cortisol concentration (9.42 ng/mL) 60 minutes after RT. This horse was observed performing stereotypical behavior and was not included in subsequent correlations with other measures. When this horse was discounted, a significant effect of sample time on cortisol concentration was found for the other 4 horses, showing a drop in cortisol concentration (Friedman test: χ^2 (6, n=4) = 13.143, $P = 0.041$). This effect was still apparent when all 5 horses were included but the last sample time (60 minutes post RT) was excluded (Friedman test: χ^2 (5, n = 5) = 13.506, $P = 0.19$). Cortisol concentration was found to be significantly lower (mean 1.29 ± 0.12 ng/mL) 15 minutes post RT than at any other sample time ($P < 0.05$). No significant effect of sample time was found for the horses showing increased cortisol concentration associated with the RT. Minimum cortisol concentrations occurred either at the initial sample time (stable: n = 2), 60 minutes post RT (n = 1), immediately (n = 1), or 5 minutes post RT (n = 1). Maximum cortisol concentrations were found at either 5 (n = 3) or 30 (n = 2) minutes post RT. The mean cortisol responses of both groups are shown in the Figure. The cortisol response of each horse is shown in Table 6. A positive correlation between baseline cortisol concentration and the age of the horse was found (rho = 0.644, n = 10, $P = 0.044$) and there was a tendency for younger horses to show a greater increase although this was not significant (rho = −0.564, n = 10, $P = 0.084$). No other factor was found to affect cortisol concentration.

Eye temperature

Static images

Eye temperatures in the stable were significantly higher (mean 32.44 ± 2.96°C) compared with those recorded immediately after WU (mean 30.06 ± 3.13°C) and immediately after RT (mean 30.56 ± 3.04°C) (Wilcoxon test: z = −2.65, $P = 0.008$; z = −2.4, $P = 0.017$, respectively). No significant difference between eye temperature recorded immediately after WU and after RT was found. There was a mean decrease in maximum eye temperature between the stable and the end of WU (−2.38 ± 1.99°C) and a slight subsequent increase after RT (+0.50 ± 1.37°C). Only horse 5 had a decrease in eye temperature (−2.75°C) between the end of WU and end of RT.

Video images

There was no difference in mean maximum eye temperature in WU and RT (video) (WU: 28.78 ± 2.32°C, RT: 28.40 ± 2.06°C). The temperatures obtained from the video footage were lower than those recorded immediately after WU and RT using static images, but a positive correlation was found between the RT video footage and static images taken immediately after WU (0 minutes pre-RT) (rho = 0.709, n = 10, $P = 0.022$) and RT video footage and static images taken immediately after RT (rho = 0.891, n = 10, $P = 0.001$). Correlation between WU video footage and static images taken immediately after WU was found, but this was not significant (rho = 0.612, n = 10, $P = 0.06$).

A significant increase in eye temperature occurred during WU (Friedman test: χ^2 [2, n = 10] = 8.60, $P = 0.014$). A significantly lower maximum eye temperature was recorded at the start (27.66 ± 2.77°C) than at after 1 minute during WU (29.65 ± 2.39°C; Wilcoxon test: z = −2.803, $P = 0.005$). No difference in eye temperature occurred during RT.

No correlation between ambient temperature and eye temperature was found apart from immediately after RT (static images) when higher ambient temperature correlated with lower eye temperatures (rho = −0.869, n = 10, $P = 0.001$). No correlation between mean maximum eye temperature (static or video) and salivary cortisol concentration was found, but the horse with the greatest increase in cortisol also had the highest eye temperature (Table 6). No effect of sex, housing, noseband, role, or age on eye temperature was found.

Association between ridden behavior and physiological measures

The only significant correlations between physiological measures and behavior related to head and neck position. There was a positive correlation between duration of low head carriage and increased salivary cortisol concentration (rho = 0.676, n = 9, $P = 0.046$). Higher eye temperatures were found to correlate with increased duration of nasal plane behind the vertical in RT both during (video: rho = 0.731, n = 10, $P = 0.016$) and immediately after RT (static: rho = 0.679, n = 10, $P = 0.031$). Lower eye temperatures were correlated with increased duration of nasal plane in front of the vertical in RT both during (video: rho = −0.806, n = 10, $P = 0.005$) and immediately after RT (static: rho = −0.879, $P = 0.001$). Although no correlation between tail swishing and cortisol or eye temperature was found, the 2 horses (horses 6 and 10) with the greatest increase in salivary cortisol spent more than 50% of the RT in this activity (Table 6). Only 1 other horse tail swished for as long.

Nonvisibility of different behaviors varied (Friedman test: χ^2 [5, n = 10] = 39.6, $P < 0.001$). Behavioral categories arranged in order of visibility (with mean ± standard deviation % duration nonvisible): tail movement (1.21 ± 2.24%); head/neck carriage (16.16 ± 4.86%); head movement (20.91 ± 3.29%); mouth (28.03 ± 5.55%), head position/angle of nasal plane (28.25 ± 9.66), and ears (56.87 ± 27.4%). Differences in visibility for all categories, apart from between mouth and head position/angle of nasal plane, were significant ($P < 0.05$).

Professional judgment

No correlation between the overall mean scores or scores for individual parameters and cortisol or mean eye temperature was

Table 6
Physiological responses, behavior (shown as % time in each) and professional assessment of individual horses in the ridden test

Horse number	1	2	3	4	5	6	7	8	9	10
Change in salivary cortisol concentration ng/mL	−2.43	−1.45	+1.24	+1.37	−1.91	+4.12	−8.11	−3.31	+3.95	+5.43
Mean maximum eye temperature °C ± SD	27.26 ± 1.11	26.26 ± 1.62	26.21 ± 0.95	29.94 ± 1.75	28.08 ± 1.47	27.22 ± 2.43	29.66 ± 2.55	28.79 ± 1.81	29.35 ± 2.48	33.41 ± 1.50
Head/neck high	5.93	50.75	0	0.95	14.7	0	0	8.16	0	0
Head/neck neutral	79.66	32.76	66.28	77.34	70.34	91.92	82.24	77.39	66.34	73.81
Head/neck low[a]	0	0	14.62	0	0	1.10	0	0	9.76	14.36
Nasal plane vertical	20.90	7.40	20.18	48.50	17.50	24.67	54.6	68.9	0.97	16.19
Nasal plane behind vertical[b1]	2.32	0	0	0.44	0	0	0.19	0	55.44	45.30
Nasal plane in front of vertical[b2]	55.82	74.29	57.09	13.31	61.34	36.63	24.50	10.04	0	1.67
Mean score[c] (veterinarians) ± SD	2.61 ± 0.57	3.29 ± 0.60	2.75 ± 0.22	3.07 ± 0.69	3.11 ± 0.30	2.36 ± 0.83	2.29 ± 0.17	2.32 ± 0.56	2.36 ± 0.25	2.32 ± 0.76
Mean score[c] (instructors) ± SD	3.14 ± 0.31	3.36 ± 0.54	2.97 ± 0.46	3.22 ± 0.34	3.25 ± 0.62	2.61 ± 1.21	2.89 ± 0.52	2.61 ± 0.58	3.58 ± 0.63	2.68 ± 0.82
Mean score[c] (riders) ± SD	3.11 ± 0.52	3.07 ± 0.34	3.04 ± 0.39	3.25 ± 0.43	3.47 ± 0.71	2.57 ± 0.51	2.54 ± 0.48	2.79 ± 0.53	2.29 ± 0.47	2.36 ± 0.50

[a] Denotes behavioral state where a significant positive correlation with changes in salivary cortisol concentration was found ($P < 0.05$).
[b] Denotes behavioral state where a significant correlation with eye temperature was found, positive ([1]) or negative ([2]) ($P < 0.05$).
[c] The higher the score, the more negative the assessment (see Table 4).

found. However, certain behavioral states were found to correlate with these scores (Table 5). The longer the duration of high head/neck carriage, the more negative the overall evaluation (rho = 0.679, n = 10, P = 0.031). When the mean scores for individual parameters (as shown in Table 4) were assessed, the following was found. Duration of head high carriage was positively correlated with lack of motivation (rho = 0.949, n = 10, $P < 0.001$), as was the duration of ears in a relaxed backward position (rho = 0.720, n = 10, P = 0.019). Low head carriage was associated with increased motivation (rho = −0.643, n = 10, P = 0.045). Horses were judged to be more confident when their head was still for a higher percentage of the time (rho = −0.644, n = 10, P = 0.044), and more energetic when their nasal plane was behind the vertical for longer (rho = −0.779, n = 10, P = 0.008). Horses were judged to be less supple the longer their nasal plane was in front of the vertical (rho = 0.756, n = 10, P = 0.011) and more resistant the longer they had their left ear forward and right ear back (rho = 0.634, n = 10, P = 0.049). A summary of behavior that was found to have positive or negative associations with physiological measures and with assessment by equine professionals is shown in Table 7.

When scoring by the different professional groups was compared positive correlations between scoring by veterinarians and instructors (rho = 0.685, n = 10, P = 0.029) and between veterinarians and riders (rho = 0.758, n = 10, P = 0.011) were found, but there was no correlation between riders and instructors. Behavior found to be associated with scores by each group differed. The veterinarians attributed poorer scores to the horses with a longer duration of high head carriage (rho = 0.64, n = 10, P = 0.046), nasal plane in front of the vertical (rho = 0.673, n = 10, P = 0.033) and ears back (rho = 0.791, n = 10, P = 0.006). The instructors attributed better scores to the horses with a longer duration of neutral head carriage (rho = −0.685, n = 10, P = 0.029), vertical nasal plane (rho = −0.648, n = 10, P = 0.043), and greater tail movement (rho = −0.673, n = 10, P = 0.033). They attributed poorer scores to horses with longer durations of ears back (rho = 0.841, n = 10, P = 0.002). The riders attributed poorer scores to the horses with a longer duration of nasal plane in front of the vertical (rho = 0.605, n = 10, P = 0.029). See Table 8 for a summary of these findings.

Table 7
A comparative summary of ridden horse behavior found to correlate with either physiological measures (cortisol and eye temperature) or professional scoring during the ridden test

	Physiological evidence	Professional assessment
Negative indicators	Low head carriage Nasal plane behind vertical	High head carriage Ears back relaxed (motivation) Left ear forward/right ear back
Positive indicators	Nasal plane in front of vertical	Nasal plane behind the vertical Low head carriage Head still

Handling test

The Cronbach's alpha value of 0.97 for reliability showed a high degree of interobserver consistency in behavior scoring during the handling tests. No significant difference in handling behavior was found in relation to sample time (stable at start/0 minutes post RT/60 minutes post RT). A significant positive correlation was found between mean maximum eye temperature (from the same sample times) and mean behavior score, with higher eye temperatures being associated with more anxious behavior (rho = 0.794, n = 10, P = 0.006). There was a positive correlation between age of horse and handling score, with younger horses showing greater anxiety (rho = −0.748, n = 10, P = 0.013). There was also a significant effect of current role on handling score (χ^2 = 6.545, df = 2, P = 0.038), with competition horses scored as being significantly more anxious than leisure horses or riding school horses ($P < 0.05$).

Discussion

The assessment made by the equestrian professionals in this study was not in accord with the physiological evidence, and there was some discrepancy between the different groups as to what behavior constituted a positive or negative indicator. All professionals (and veterinarians in particular) deemed a high head carriage and the nasal plane being in front of the vertical to be negative signs. However, the instructors scored neutral head carriage and vertical nasal plane as positive indicators. This lack of agreement between equestrian professionals is likely to be the result of assessing ridden horses for different purposes. Although instructors and riders will generally be looking to improve ridden behavior, veterinarians will be assessing soundness and other health parameters. In the study by König von Borstel et al. (2011) in which ride-ability was judged to be better when the horse predominantly carried its nose behind the vertical, the assessment was carried out by riders rather than by judges on the ground. The latter study was conducted in Germany and may reflect international differences in equestrianism, However, the riders in the current study also appeared to consider the position of the nasal plane in front of the vertical to be a negative sign, even when observing the horse from "the ground." These findings provide further evidence of departure from the guidelines provided by the FEI stating that the

Table 8
Behavioral states associated with positive and negative evaluation by different equestrian professionals

	Veterinarians	Instructors	Riders
Negative indicators	High head carriage* Nasal plane in front of vertical* Ears back**	Ears back**	Nasal plane in front of vertical*
Positive indicators		Neutral head carriage* Vertical nasal plane* Increased tail movement*	

*$P < 0.05$, **$P < 0.01$.

nasal plane should at all times be positioned in front of the vertical (FEI, 2009).

The cortisol response to this RT varied with horses showing peaks that appeared to relate specifically to the ridden work at varying times. As noted previously, cortisol levels vary with training and level of fitness as well as the physical demands of the activity (Marlin and Nankervis, 2002; Fazio et al., 2006). The positive association found between increased cortisol and duration of low head carriage indicates that this position is more physically demanding than neutral or high head carriage. This effect is comparable to the increases in eye temperature found in horses lunged in a training aid (Pessoa) (Hall et al., 2011). The horse with the greatest increase in cortisol (that appeared to relate to the RT) was "novice eventing fit" at the time of the study. This response may be indicative of anxiety, as shown by Schmidt et al. (2010). All 5 horses that had decreases in cortisol had minimum values at 15 minutes after the RT, suggesting increased relaxation. The positive effect of exercise on stabled horses has been demonstrated in the past (Freire et al., 2009), but in the current study no horse was stabled continuously. We found a positive correlation between age and baseline cortisol concentration but no link between cortisol and housing or current ridden role. It is possible that salivary cortisol analysis can provide an insight into the link between ridden behavior and physical and/or mental stress but fitness and experience must first be taken into account.

More immediate measures of stress can be made using surface temperature changes associated with redirected blood flow (Stewart et al., 2007). Increases in eye temperature have previously been shown to relate to stress responses in ridden horses (Valera et al., 2012). In the current study, the horse with the greatest cortisol increase also had the highest eye temperature, which, unlike the other horses tested, increased in the ridden work compared with the stable. In other species, anxiety/fear responses have been associated with surface temperature changes of $<0.2°C$ (humans: Ekman et al., 1983; rhesus monkeys: Nakayama et al., 2005). The eye area of the horse is surrounded by hair, which may affect the temperature readings. Also, the thermal camera used in this study was not sensitive enough to reliably record such subtle changes in temperature; further work is needed to confirm the precise nature of surface temperature changes in the horse before conclusions can be confidently drawn.

Correlations between specific behaviors and physiological measures provide some insight into the effect of aspects of ridden work. Increased eye temperature was found to correlate with increased time spent with the head being carried with the angle of the nasal plane behind the vertical. This correlation could have been an artifact of the angle at which the video thermal image was taken but was also found with eye temperatures taken from static images immediately after the ridden work. Carrying the head at an angle behind the vertical could relate to increased anxiety and/or reduced vision (McGreevy et al., 2010). The related increases in eye temperature provide further evidence of the potential negative effect of this particular head carriage in the ridden horse. The eye temperatures gathered using static images were higher than those obtained from a greater distance using video, but correlation between the 2 indicates that although actual temperatures may not be accurate, relative values were consistent. Validation of this technology is required before it can be reliably used to interpret behavior but further investigation is warranted.

The relative visibility of certain aspects of ridden behavior varied with tail movement being most visible. Agitation of the tail is identified by the FEI (2009) as a sign of nervousness, tension, or resistance; the 2 horses that had RT-related cortisol increases spent more than 50% of the time tail-swishing. Other more subtle signs such as grinding the teeth (FEI, 2009) were less visible. More conclusive associations may be apparent if a continuous more comprehensive behavioral recording system could be developed. Further identification of behavioral signs based on objective measures is required to ensure that consistent and accurate judgment of ridden horse behavior can be made.

A positive correlation between eye temperature and anxiety-related handling behavior was found. In this part of the study, the observers were provided with guidance on what behavior to look for when scoring each horse and this resulted in a high degree of interobserver agreement. The guidelines provided by the FEI (2009) on aspects of ridden horse behavior that are positive or negative indictors were at least in part substantiated by the physiological measures taken in the current study. No specific guidelines were provided for the ridden horse assessment by equestrian professionals, and only the instructors were found to make judgments that agreed in part with FEI guidelines (FEI, 2009) (Table 8). The distinction made between tail-swishing and tail-swinging was not always clear and although this behavior was found to be most visible, it is still open to differences in interpretation.

Inconsistency between the scoring of the ridden behavior by the equestrian professionals may well have been the consequence of interpreting behavioral signs differently and the need for comprehensive, evidence-based guidelines for what ridden behavior is indicative of "good" or "poor" performance is clear. Such guidelines would allow the early identification of potential problems. Ridden behavior problems are highly prevalent, particularly in those animals used primarily for leisure purposes (Hockenhull and Creighton, 2013) and consequent wastage/euthanasia of horses could be minimized if warning signs were identified more accurately and responded to earlier. Poor ridden performance can result from underlying, often subclinical, health problems (Mitchell, 2001) and stress in ridden work can increase the risk of health issues such as equine gastric ulcer syndrome (McClure et al., 2005). We are unlikely to achieve universal consensus on how to assess ridden horse behavior. However, we must continue to acquire further evidence to ensure consistent and valid interpretation of behavioral signs in ridden horses.

Conclusions

Behavioral signs and physiological measures need further evaluation to ensure the accurate interpretation of ridden horse behavior and consequently reduce the occurrence of stress-related health and behavioral problems.

Acknowledgments

We thank the riders who took part in this study and the equestrian professionals who scored the ridden horse behavior, in particular the veterinarians from the Minster Veterinary Centre

(Nottinghamshire, UK) and from the School of Veterinary Science and Medicine, University of Nottingham, UK.

Conflict of interest: None.

References

Cook, N.J., Schaefer, A.L., Church, J.S., 2006. Nutritional therapy modulates stress responses of elk (*Cervus elaphus Canadensis*) to removal of velvet antler. Online J. Vet. Res. 10, 20–25.

Ekman, P., Levenson, R.W., Friesen, W.V., 1983. Autonomic nervous system activity distinguishes between emotions. Science 221, 1208–1210.

Fazio, E., Calabrò, G., Medica, P., Messineo, C., Ferlazzo, A., 2006. Serum cortisol levels of quarter horses: circadian variations and effects of training and western riding events. In: Lindner, A. (Ed.), Management of Lameness Causes in Sport Horses. Conference on Equine Sports Medicine and Science. Wageningen Academic Publishers, Cambridge, U.K, pp. 175–179.

Fazio, E., Molinari, P., Medica, P., Messineo, C.,, Ferlazzo, A., 2008. Circulating cortisol levels of Arab sport horses before and after gymkhana riding events: effects of training state, gender and age. In: Lindner, A. (Ed.), The Acute Poorly Performing Sport Horse. Conference on Equine Sports Medicine and Science. Wageningen Academic Publishers, Cambridge, U.K, pp. 165–171.

Fédération Equestre Internationale, 2009. FEI Rules for Dressage Events, 23rd ed. Fédération Equestre Internationale, Lausanne, Switzerland.

Freire, R., Buckley, P., Cooper, J.J., 2009. Effects of different forms of exercise on post inhibitory rebound and unwanted behaviour in stabled horses. Equine Vet. J. 41, 487–492.

German National Equestrian Federation, 1997. The Principles of Riding: The Official Instruction Handbook of the German National Equestrian Federation. German National Equestrian Federation, Kenilworth, U.K.

Hall, C., Burton, K., Maycock, E., Wragg, E., 2011. A preliminary study into the use of infrared thermography as a means of assessing the horse's response to different training methods. J. Vet. Behav.: Clin. Appl. Res. 6, 291–292.

Hall, C., Huws, N., White, C., Taylor, E., Owen, H., McGreevy, P.D., 2012. Assessment of ridden horse behavior. J. Vet. Behav.: Clin. Appl. Res. 8, 62–73.

Heleski, C.R., McGreevy, P.D., Kaiser, L.J., Lavagnino, M., Tans, E., Bello, N., Clayton, H.M., 2009. Effects on behaviour and rein tension on horses ridden with or without martingales and rein inserts. Vet. J. 181, 56–62.

Hockenhull, J., Creighton, E., 2013. The use of equipment and training practices and the prevalence of owner-reported ridden behaviour problems in UK leisure horses. Equine Vet. J. 45, 15–19.

Kaiser, L., Heleski, C.R., Siegford, J., Smith, K.A., 2006. Stress-related behaviors among horses used in a therapeutic riding program. J. Am. Vet. Med. Assoc. 228, 39–45.

König von Borstel, U., Pasing, S., Gauly, M., 2011. Towards a more objective assessment of equine personality using behavioural and physiological observations from performance test training. Appl. Animal. Behav. Sci. 135, 277–285.

Marlin, D., Nankervis, K., 2002. Equine Exercise Physiology. Blackwell Science, Oxford, U.K, pp. 130–132.

McClure, S.R., Murray, M.J., Carithers, D., Gross, S.J., Holste, J.E., 2005. Gastric ulceration in horses exposed to training and activities typical for recreational showing. Am. Assoc. Equine Practitioners.. Proceedings 51.

McGreevy, P.D., Harman, A., McLean, A., Hawson, L., 2010. Over-flexing the horse's neck: a modern equestrian obsession? J. Vet. Behav.: Clin. Appl. Res. 5, 180–186.

Mitchell, R.D., 2001. Prevalence of gastric ulcers in hunter/jumper and dressage horses evaluated for poor performance. Proceedings of the 21st Annual Meeting of the Association of Equine Sports Medicine 2001, 74-77.

Nakayama, K., Got, S., Karaoke, K., Nakamura, K., 2005. Decrease in nasal temperature of rhesus monkeys *(Macaca mulatto)* in negative emotional state. Physiol. Behav. 84, 783–790.

Ödberg, F.O., 1987. Chronic stress in riding horses. Equine Vet. J. 19, 268–269.

Schmidt, A., Aurich, J., Mőstl, E., Müller, J., Aurich, C., 2010. Changes in cortisol release and heart rate and heart rate variability during the initial training of 3-year-old sport horses. Hormones Behav. 58, 628–636.

Stewart, M., Webster, J.R., Schaefer, A.L., Cook, N.J., Scott, S.L., 2005. Infrared thermography as a non-invasive tool to study animal welfare. Animal Welfare 14, 319–325.

Stewart, M., Webster, J., Verkerk, G., Colyn, J., Schaefer, A., 2007. Non-invasive measurement of stress in dairy cows using infrared thermography. Physiol Behav. 92, 520–525.

Valera, M., Bartolomé, E., Sánchez, M.J., Molina, A., Cook, N., Schaefer, A., 2011. Changes in eye temperature and stress assessment in horses during show jumping competitions. J. Equine Vet. Sci. 32, 827–830.

Van Breda, E., 2006. A nonnatural head-neck position (*Rollkur*) during training results in less acute stress in elite, trained, dressage horses. J. Appl. Animal Welfare Sci. 9, 59–64.

Von Borstel, U.U., Duncan, I.J.H., Shoveller, A.K., Merkies, K., Keeling, L.J., Millman, S.T., 2009. Impact of riding in a coercively obtained Rollkur position on welfare and fear of performance horses. Appl. Animal Behav. Sci. 116, 228–236.

Weeks, J., 1996. Equine agitation behaviours. Equine Practice 18, 23–24.

Yarnell, K., Hall, C., Billett, E., 2013. An assessment of the aversive nature of an animal management procedure (clipping) using behavioral and physiological measures. Physiol Behav. 118, 32–39.