Dually Investigated: the effect of a pressure headcollar on the behaviour, discomfort and stress of trained horses

Carrie Ijichi[1*], Hayley Wild[1], Francesca Dai[2], Alexandre Bordin[1], Heather Cameron-Whytock[1], Samuel J. White[1], Kelly Yarnell[1], Gareth Starbuck[1], Aurelie Jolivald[1], Lauren Birkbeck[1], Sarah Hallam[1] & Emanuela Dalla Costa[2]

[1] School of Animal, Rural & Environmental Science, Nottingham Trent University, Brackenhurst Campus, UK

[2] Università degli Studi di Milano, Department of Veterinary Medicine, Italy

**carrie.ijichi@ntu.ac.uk*

Abstract

The Dually™ is a control headcollar designed to improve equine behaviour during handling challenges by applying greater pressure than a standard headcollar. Previous research indicated it did not improve compliance in naïve horses but did result in higher Horse Grimace Scale scores (HGS) indicative of discomfort. However, subjects had not been trained to step forward to release the pressure applied by the headcollar. The current study aimed to determine the effect of training on behaviour and physiology of horses wearing the Dually™ headcollar during handling challenges. To this end, subjects received three training sessions prior to completing two handling tests in which they crossed distinct novel obstacles, one wearing a Dually™ with a line attached to the pressure mechanism and one attached to the standard ring as a control. Behaviour was coded by hypothesis blind researchers: time to cross the obstacle and proactive refusal (moving away from the obstacle) were recorded as indicators of compliance and the Horse Grimace Scale was used to measure discomfort caused by each configuration of the device. Infrared thermography of ocular temperature, heart rate variability (RMSSD and low/high frequency ratios (LF/HF)) and salivary cortisol were measured as indicators of arousal. Data from the previous study on Naïve horses was also included to compare responses to the Dually in Naïve and Trained horses. Training resulted in a decrease in RMSSD (p = 0.002) and an increase in LF/HF (p=0.012), compared to rest, indicating arousal. As per the original study, horses did not complete the tests more quickly in the Dually, compared to control (p=0.698). Trained horses from this study tended to be more proactive in the Dually compared to Controls (p=0.066) and significantly more so than Naïve horses from the previous study (p=0.002) suggesting that behaviour becomes less desirable during early Dually training. Yet, stress and HGS indicators were not higher in the Dually compared to Control during testing. Results suggest the Dually has a negative effect on behaviour but not on stress or discomfort during short handling challenges. Further research is warranted to determine the long-term effect of Dually experience on behaviour and welfare.

Keywords: heart rate variability; infrared thermography; salivary cortisol; horse grimace scale; proactivity; horse welfare

1. Introduction

The horse is a large prey animal for which domestication has dampened, but not extinguished, innate biological flight responses (Brubaker and Udell, 2016). These responses make it difficult to retain stimulus control at all times (McGreevy and McLean, 2007) as environmental stimuli often exert more control over the horse's behaviour than their human handler is able to. Williams and Ashby (1995) state 20% of accidents occur during handling and allude to horse behaviour being the primary cause. Similarly, Sandiford et al., (2013) reported 12% of patients admitted to a UK hospital with horse related injuries sustained them in non-ridden accidents. Therefore, it is understandable that many owners seek solutions to reduce such risky behaviour during daily interactions, often by using devices which increase the salience of human cues in order to compete with environmental stimuli.

The Dually™ headcollar is a widely used, commercially available control headcollar which increases the pressure a handler can apply in order to maintain control of a horse. It therefore works using negative reinforcement: pressure from the headcollar should release when the horse offers the desired response. It has two settings: a standard ring under the chin and two side rings which operate an inbuilt pressure-release mechanism. When the lead-rope is attached to the side ring, if the horse pulls back or fails to walk forward when pressure is applied by the handler, the inbuilt mechanism tightens, increasing the level of pressure exerted around the jaw and nose of the horse (Roberts, 1999). The patent for the Dually™ states *“It is extremely effective for training the animal to lead, to stand still, to walk into a truck or trailer, to walk slowly through narrow passages, to walk over unfamiliar objects…”* (Roberts, 1999). However, research investigating bridles which apply pressure to similar sensitive facial structures highlights welfare concerns (Doherty et al., 2017; Fenner et al., 2016; McGreevy et al., 2012). Further, Ijichi et al., (2018) found the Dually™ did not improve compliance in naïve horses but did result in higher Horse Grimace Scale scores (HGS). However, in the previous study, subjects were naïve to the Dually™ and had not been trained to give the desired response, resulting in release of pressure. Therefore, the headcollar may still be valuable in modifying the behaviour of horses that are trained to offer the desired response to release the pressure.

The aim of the current study was to determine the effect of the Dually™ on behaviour and physiology of trained horses during handling challenges. To this end, subjects received three training sessions prior to completing two novel handling tests, one wearing a Dually™ with a line attached to the pressure mechanism and one attached to the standard ring as a control. Each test consisted of crossing two different novel objects to avoid habituation. Time to cross the obstacle and proactive refusal (moving away from the obstacle) were blind scored as indicators of compliance (Ijichi et al., 2013). The Horse Grimace Scale was scored by an observer blind to the experimental study design (Dalla Costa et al., 2014). Ocular temperature measured by infrared thermography (IRT) (Yarnell et al., 2013), heart rate variability

(HRV) (von Borell et al., 2007) and salivary cortisol (Hughes et al., 2010) were measured as indicators of stress and arousal. Data from the previous study on naïve horses completing similar tests (Ijichi et al., 2018) was also included to compare the responses of trained and naïve horses. Results were compared between Control and Dually™ in Trained horses and between Naïve and Trained horses. It was hypothesised that Dually™ Training would result in improved compliance, and reduced arousal and HGS scores compared to Naïve Dually™ horses and improved compliance compared to Trained Control.

2. Method

A sample of 16 resident Nottingham Trent University mixed-breed horses (10 geldings and 6 mares) aged between 4 and 22 years (mean = 13 years ± 4.85) participated in the study. Subjects were housed and managed as per normal protocol. In general, horses were provided with forage three times a day, concentrate feed dependent on workload and nutritional requirements and had access to fresh water at all times. At the time of testing, subjects were housed individually or with a companion during the day and turned out at night. The study took place in an enclosed outdoor research arena at Brackenhurst campus between 14th and 17th May 2019. Horses were paired according to companion preference and both were present in their allocated pair in the arena during training and testing to prevent isolation stress. All horses were handled by the same experimental handler for all training and testing sessions (CI).

2.1 Training Protocol

Subjects underwent three 10-minute training sessions wearing a correctly fitted Dually™ headcollar (Roberts, 1999) with the lead-rope attached to the left side ring. All three training sessions were carried out on the same day over a 1-hour period for each pair, alternating 10-minute training sessions with 10 minutes of rest. Each training session focussed on specific desired responses: stop and step-forward; accelerate and decelerate; Stop, step forward, accelerate, decelerate, back-up two steps (Table 1). Pair order was pseudo-randomised to account for subject availability. A training chute was marked along the short side of the arena using standard jump poles laid end-to-end along the ground. These poles were placed 2m out from the fence and the total length was 12m. This area was filmed using a Canon Legria HFR606 camcorder.

The handler held the lead-rope approximately 2 inches from the side ring and maintained a light contact. Horses were led to the training chute and given a cue to offer the relevant response every four strides by applying forward or backward pressure to the lead-rope. Pressure increased until the desired response was offered and then immediately released. No vocal or other tactile stimuli were used. Once at the end of the chute, the handler released the contact, scratched the horse on the withers and offered verbal praise in a soft tone. They allowed the horse to lower their head if they chose and walk at their preferred speed as they guided them in an arc around to the start of the training chute. Once at the start of the chute this process was repeated until the 10-minute training session was complete, whereupon the horse was led to the rest area. This training protocol resulted in a high number of trials (Table 1) with inter-trial intervals of approximately 5 seconds, but regular short breaks of approximately 30 seconds every three-four trials and larger 10-minute breaks between sessions to consolidate learning and minimise arousal. After completing three training sessions, subjects were returned to their stables. All subjects were able to stop, step forward, accelerate, decelerate and back-up two steps at the end of the training day (Table 1). Subjects had a rest day following training with testing on the subsequent day.

Table 1. Targeted responses and number of trials per session and in total (per horse).

	Task	Number of Trials
Training Session 1	Stop & step forward	Mean = 61 (±13)
Training Session 2	Accelerate & decelerate	Mean = 38 (±9)
Training Session 3	Stop, step forward, accelerate, decelerate, back-up two steps	Mean = 58 (±12)
	Total	**Mean = 157 (±22)**

2.2 Testing Protocol

2.2.1 Novel Handling Tests

For the novel handling test, subjects were asked to cross two distinct obstacles (Test A & B) to avoid habituation from the first attempt. Subjects completed one test with a lunge-line attached to the side ring (Dually™) and one attached to the under-chin ring (Control) as per Ijichi et al. (2018). Test and treatment order were randomised in a counterbalanced design. Test A consisted of a 2.5m x 3m yellow tarpaulin secured to the ground by tent pegs; a piece of red carpet was placed on top of the tarpaulin allowing for a trim of approximately 0.75m of tarpaulin to be visible. Test B consisted of a green camouflage tarpaulin secured to the ground with individual tent pegs with a piece of pale blue carpet placed on top of the tarpaulin to leave a trim visible as per Test A.

The start of each test was marked by a single horizontal pole placed on the ground 2m in front of the obstacle. The handler walked the horse toward the obstacle and asked the horse to cross by applying pressure to the headcollar with no additional aids, verbal commands or further encouragement, as per the training sessions. Pressure was applied if the horse stopped, moved sideways or away from the obstacle and was immediately released when the horse took a step toward the obstacle in accordance with learning theory (McGreevy and McLean, 2007). During treatment, this meant the in-built pressure mechanism activated around the nose of the horse, whilst under control only poll pressure occurred.

2.2.2 Behaviour Analysis

The area covering the pole and the tarpaulin was filmed using Canon Legria HFR606 for retrospective analysis of behaviour by a hypothesis blind researcher (AB). Crossing time for each test began when the subject’s front hoof crossed over the pole and bore weight on the ground. Time stopped when the last rear hoof bore weight on the tarpaulin. Horses engage their rear legs first when transforming into faster gaits. Therefore, horses that showed a flight response on the tarpaulin were not given faster crossing times. For the attempt to be classed as a successful crossing, all four hooves must have been placed onto the tarpaulin, which excludes those where the horse completely or partially jumped over the obstacle. Incomplete crossings resulted in the horse being returned to make another attempt. A time limit of 3 minutes was allotted for each attempt as previous research indicated that subjects which had not completed the test within this time were unlikely to do so (Ijichi et al., 2013). Once the 3-minute threshold had been reached the test was ended. A crossing time of 180 seconds was given to any horse reaching this time limit.

Refusal behaviour was defined as any behaviour which did not contribute to crossing the object (Ijichi et al., 2013): moving backwards, sideways, forwards but away from the tarpaulin, rearing or remaining stationary. Refusal that lasted for 10 seconds or more was analysed to determine how proactive that refusal was. Nine horses refused both tests for 10 seconds or more, providing data for paired statistical tests. Proactive refusal was defined as any refusal behaviour that involved movement: moving backwards, sideways, forwards but away from the tarpaulin or rearing.Proactive refusal was then recorded as the percent of total refusal time for any individual which showed refusal behaviour (which included remaining stationary) and reported as “proactive behaviour”. A higher value indicated a greater amount of proactive behaviour as per Ijichi et al. (2013). This behaviour is of interest because proactive behaviour is more typically associated with horse-handler accidents since refusing to move does not involve unpredictable behaviour.

2.2.3 Salivary Cortisol

Baseline saliva samples were taken from subjects immediately prior to each Training and Testing session and then samples were taken again 10 minutes after to allow any cortisol changes to reach the saliva (Yarnell et al., 2013). Baseline salivary cortisol measures were not taken in the stable at the same time as heart rate variability (see section 2.2.5) as cortisol fluctuates with diurnal rhythms (Hoffis et al., 1970). Therefore, if baseline cortisol was taken on a prior date and time of day, changes from baseline may be the result of confounding factors, rather than experimental conditions per se. By taking baseline saliva samples immediately prior to Training and Testing and calculating changes in salivary cortisol (rather than using the absolute concentrations), diurnal fluctuations cannot impact the results.

Saliva samples were taken with an Equisal swab gently moved over the tongue and lips of the subject (Ijichi et al., 2019). These swabs are specifically designed for use in horses and are routinely used to test for tapeworm. Subjects were familiar with similar sampling as they are regularly wormed, tested for worms and have saliva taken for cortisol analysis for other studies. Samples were placed in a cooler box with ice packs before being transferred to the laboratory freezer within 2 hours of collection.

A competitive ELISA (Cortisol ELISA, IBL International, Hamburg, Germany) developed for quantitative analysis of free cortisol in human saliva was used. This Elisa has been validated and used in horses repeatedly (e.g. Sauer et al., 2020; Scheidegger et al., 2016)The assay was performed according to manufacturer instructions. Saliva samples were thawed and centrifuged at 500 rpm at room temperature for 3 min using Hereaus Fresco 17 centrifuge (ThermoScientific, West Sussex, United Kingdom). The plate was shaken for 5 min using an orbital shaker (Flow Laboratories DSG Titertek, Pforzheim, Germany). The plate was washed 4 times with 1X wash buffer by gently squirting the buffer into each well with a squirt bottle. Optical density was measured by a Multiscan EX (Thermo Labsystems, Vantaa, Finland). The results were calculated using four-parameter-logistic as recommended by the manufacturer. To determine the effect of training, the average of the three sessions was calculated. The change in salivary cortisol from pre-test to post-test A and B were used to determine the difference between Dually and Control, to account for diurnal fluctuations in cortisol (Hoffis et al., 1970).

2.2.4 Infrared Thermography

A FLIR E4 thermal imaging camera (FLIR Systems, USA.) was used to record eye temperature (°C). Baseline IRT images were taken immediately before each Training and Testing session and subsequent samples taken immediately after. Baseline IRT was not taken in the stable at the same time as heart rate variability (see section 2.2.5) as this fluctuates with environmental conditions (Church et al., 2014). Therefore, changes from baseline may be the result of confounding factors, rather than experimental conditions per se. After pre-session saliva samples were collected, horses were led to the measurement chute. This consisted of two jump poles laid parallel 1m apart. A small cavaletti block at one end marked of the chute to mark where the horses head should be once stationary. Two cavaletti were positioned 1m away from this central marker on either side of the measurement chute to mark where the IRT camera should be positioned when taking images of the left and right eye (90° from the eye and 1m away (Ijichi et al., 2020)). This kept the horse straight and in the same direction for all images and standardised the optimal camera angle and distance as the angle of measurement significantly affects temperature readings (Ijichi et al., 2020). Subjects had experienced this measurement chute and its constituent parts before but had not been systematically habituated to it. However, whilst this might have caused them slight arousal, it would not explain any differences before and after training or between Dually and Control readings.

Images were analysed using FLIR Tools software (ver. 5.9.16284.1001) to obtain a measurement for each eye. All images were analysed by the same two researchers (C.I. & H.W.). Eye temperature recordings were the maximum temperature within the palpebral fissure from the lateral commissure to the lacrimal caruncle (Yarnell et al., 2013). A mean of the left and right eyes was calculated for each subject, pre and post-test, for each training session and test. The average temperature change was calculated to determine the effects of training. The change in average temperature from pre-test to post-test was used to account for individual differences and fluctuations in core temperature due to changing environmental conditions.

2.2.5 *Heart Rate Variability*

Heart rate variability was recorded with a Polar Equine V800 portable heart rate monitor for baseline and all Training and Testing sessions (Polar Electro Oy, Kempele, Finland). Some authors have recently questioned whether Polar devices measure HRV as accurately as Electroencephalogram (ECG) devices (Pearson et al., 2019). However, such devices are much less readily available and Polar monitors are commonly used in research and have also been argued to be valid in horses (McDuffee et al., 2019). Therefore, they were deemed appropriate for the current study as they were used in conjunction with other measures of stress.

The surcingle was fitted to each subject after the first saliva collection at the start of Training and Testing days and remained on until the subject had completed data collection for the day. The girth area of each subject was wetted to ensure contact and enhance electrical conductivity. Electrodes were positioned in the region of the upper left thorax and the ventral midline (Yarnell et al., 2013). The receiving watch was looped onto the surcingle to ensure it remained within connectivity boundaries at all times.

Baseline heart rate variability was recorded to determine changes as a result of training and testing. To mitigate any potential impact of anticipatory stress, baseline heart rate and heart rate variability parameters were recorded after a period of wearing the heart rate monitor undisturbed in the home stable. Data was collected between 10.30am and 3.30pm between 11th – 14th February 2019. Horses were loosely tethered in their home environment with a headcollar and leadrope and fitted with a Polar Equine V800 Science heart rate monitor before being released. RR interval data was recorded continuously for 35 minutes while the horses were left undisturbed in their home environment. Potential environmental disturbances were recorded by an observer. Thereafter, horses were caught and tethered again, the recording stopped and the heart rate monitor removed. If no environmental disturbance was observed during the recording, mean heart rate and heart rate variability readings were extracted from the section of the recording between 25 and 30 minutes. If an environmental disturbance was observed that visibly affected heart rate (n=2: neighbouring horse removed), readings were taken from the 5 minutes immediately preceding that disturbance.

For Training and Testing, subjects were allowed 5 minutes to habituate to the surcingle, deemed to be sufficient as all subjects have previously worn these heart monitors on several occasions. Heart rate recording commenced when the horse left the measurement chute to begin testing and ceased when the horse re-entered the measurement chute post-test after the last training or testing session of the day.

Kubios software (version 3.0.2 Biomedical Signal Analysis and Medical Imaging Group, Department of Applied Physics, University of Eastern Finland, Kuopio, Finland) was used to analyse heart rate data and determine HRV. Artefact correction was set to custom level 0.03, removing RR intervals varying more than 30% from the previous interval. Trend components were adjusted using the concept of smoothness priors set at 500ms, to avoid the effect of outlying intervals (Ille et al., 2014).

Frequency Domain analysis was set at >0.01 - ≤0.07 for Low Frequency (LF) and > 0.07 - ≤0.5 for High Frequency (HF) (Stucke et al., 2015). The full recording from leaving the IRT measurement chute to returning after completing each training or test session was selected for analysis. RMSSD values were used as these reflect high frequency beat-to-beat variations indicative of vagal activity (Stucke et al., 2015). In addition, Frequency Domain Analysis (FDA) was conducted using a fast Fourier transformation which were expressed as ratios for enhanced comparability (Stucke et al., 2015). The ratio of Low to High Frequency (LF/HF) reflects both parasympathetic and sympathetic tone as well as cardiac sympatho-vagal balance. The average RMSSD and LF/HF for the three training sessions was calculated to determine the effects of training.

2.2.6 Horse Grimace Scale

During testing, images were taken of each subject with a Panasonic camera (Model, DMC-FZ72, Japan). The photographer (H.W.) used a zoom lens to take detailed images of the subject’s face from a distance of approximately 3m. Images were included in analysis if the lunge line formed a straight line from the handler’s hand to the ring of the headcollar, indicating that pressure was being applied to the headcollar in that instance. Therefore, subjects who completed the task without hesitation did not provide images for analysis, as no pressure was required to indicate they should walk forward. Crossing time also influenced the number of images available for each subject. Images that were clearly in focus were preferentially selected. A total of 256 photographs (Control: subjects with images = 12, mean images per subject = 8.67; Dually: subjects with images = 12, mean images per subject = 10) were then analysed against the Horse Grimace Scale (Dalla Costa et al., 2014) by a researcher blind to the research hypothesis (FD).

Where an area of the face (facial action unit) was obscured it was not scored. The mean score for each Facial Action Unit from all images was calculated and then totalled to give the HGS score for each subject in each treatment.

2.2.7 Retrospective Analysis

To determine a potential effect of training on behaviour and physiology in horses wearing a Dually™ headcollar, previously collected data from 20 naïve horses who had not been trained in a Dually™ headcollar was also included (Ijichi et al., 2018). These subjects underwent the same testing procedure over novel objects, full details of which are reported by Ijichi et al (2018). Eye temperatures, crossing times and proactive behaviour were available for these subjects, but not HRV or salivary cortisol. Images of the subject's faces were re-analysed by the same researcher (FD) using the method stated in 2.2.6 in order to provide comparable data. A total of 150 images was available for analysis (Control: subjects with images= 13, mean images per subject = 6.5; Dually: subjects with images = 12, mean images per subject = 7.5). The behaviour, HGS and IRT of Trained and Naïve horses was then compared.

2.3 Ethics

The yard manager provided informed consent for all subjects via the completion of a participant information form. Both researchers and the manager had the right to withdraw a subject at any time, for any reason, until the point of data analysis. Prior to commencement, the current study was authorised by the Nottingham Trent University Ethics Committee.

2.4 Statistical Analysis

Statistical analysis was carried out using RStudio (RStudio Development Team, 2020). Shapiro-Wilks tests were used to test the distribution of the residuals between paired variables. Differences between baseline or pre-training and post-training physiology, pre and post-testing, and between Control and Dually™ treatments were investigated using either Paired T-tests or Wilcoxon tests as appropriate for normality. Shapiro-Wilks tests were used to test the distribution of variables and Levene Tests were used to test homogeneity of variance for independent tests of difference. Differences between Naïve and Trained horses were tested using Independent T-tests or Mann Whitney U-tests as appropriate for normality and homogeneity of variance. Tests of difference between Trained and Naïve were only conducted if there was no difference in Control. Otherwise, differences observed may have been due to different samples. Post-hoc effect sizes were then calculated as per Field et al. (2012).

3. Results

3.1 Effect of Training on physiology

RMSSD was significantly lower on average during training, compared to baseline (Paired T-test: T = -3.98, N = 12, P = 0.002, D = 0.754). LF/HF was significantly higher on average during training, compared to baseline (Wilcoxon: V = 78, N = 14, P = 0.021, D = -0.541). No other indicators of stress were significantly different between rest and training (Table 2).

Table 2. Differences in physiology as a result of training. Paired T-Tests (PTT) and Wilcoxon tests (W) are used as appropriate for normality.

Variable	**Treatment**	**Mean/ Median**	**SD/ IQR**	**Test**	**V/T**	**P**	**Effect Size**	**N**
IRT Change (ºC)	Pre-Training	35.9	0.91	PTT	0.79	0.441	0.207	15
	Post-Training	36.9	0.52					
RMSSD (ms)	Baseline	103.64	44	PTT	-3.98	**0.002**	**0.754**	12
	Training	49.15	16.21					
LF/HF	Baseline	0.87	0.6	W	78.00	**0.021**	**-0.541**	14
	Training	1.18	0.73					
Cortisol (µg/dL)	Pre-Training	0.61	0.46	W	39.00	0.144	-0.365	16
	Post-Training	0.48	0.58					

3.2 Effect of Testing on physiology

RMSSD was significantly lower after testing for both Dually™ (Paired T-test: T = 3.23, N = 12, P = 0.007, D = 0.667) and Control (Wilcoxon: V = 102, N = 12, P < 0.001, D = 0.989). There was a tendency for LF/HF to increase after both Dually™ (Paired T-test: T = -1.81, N = 14, P = 0.094, D = 0.448) and Control (Wilcoxon: V = 23, N = 14, P = 0.067, D = -0.916). No other variables differed following Testing (Table 3).

Table 3. Differences in physiology as a result of Testing. Paired T-Tests (PTT) and Wilcoxon tests (W) are used as appropriate for normality.

Variable	Treatment	Mean/Median	SD/IQR	Test	V/T	P	Effect Size	N
IRT (ºC)	Pre-Dually	35.74	±0.92	PTT	0.30	0.765	0.078	16
	Post-Dually	35.68	±1.05					
	Pre-Control	35.6	±0.8	PTT	0.34	0.741	0.087	
	Post-Control	35.54	±0.75					
RMSSD (ms)	Baseline	103.64	±43.9	PTT	3.23	**0.007**	**0.667**	12
	Post-Dually	48.34	±26.64					
	Baseline	87.43	65.23	W	102.00	**<0.001**	**-0.989**	
	Post-Control	49.57	24.03					
LF/HF	Baseline	0.87	±1.02	PTT	-1.81	**0.094**	**0.448**	14
	Post-Dually	2.54	±2.71					
	Baseline	0.56	0.6	W	23.00	**0.068**	**-0.916**	
	Post-Control	1.5	1.53					
Cortisol (µg/dL)	Pre-Dually	0.33	0.66	W	57.00	0.587	-0.136	16
	Post-Dually	0.28	0.26					
	Pre-Control	0.33	0.66	W	46.00	0.274	-0.273	
	Post-Control	0.29	0.33					

3.3 Differences between Treatment and Control

Proactive behaviour had a tendency to be significantly higher in the Dually™, compared to the Control (Paired T-Test: T = 2.214, N = 9, P = 0.066, D = 0.6). No other differences were observed between Treatment and Control (Table 4).

Table 4. Differences in behaviour and physiology between Dually and Control in Trained horses. Paired T-Tests (PTT) and Wilcoxon tests (W) are used as appropriate for normality.

Variable	Treatment	Mean/ Median	SD/IQR	Test	V/T	P	Effect Size	N
HGS	Dually	1.99	±0.75	PTT	-1.22	0.247	0.345	12
	Control	1.7	±0.93					
IRT Change (ºC)	Dually	-0.06	±0.67	PTT	0.023	0.982	0.008	16
	Control	-0.06	±0.82					
RMSSD (ms)	Dually	49.57	±24.03	PTT	0.206	0.840	0.053	16
	Control	48.34	±26.64					
LF/HF	Dually	1.91	1.95	W	81	0.528	-0.158	16
	Control	1.5	1.53					
Cortisol Change (µg/dL)	Dually	-0.001	0.3	W	69	0.980	-0.006	16
	Control	-0.002	0.3					
Crossing Time (secs)	Dually	23.3	57.5	W	76	0.698	-0.097	16
	Control	20.7	47.75					
% Proactivity	Dually	53.29	±26.12	PTT	2.124	**0.066**	**0.600**	9
	Control	30.17	±36.77					

3.4 Differences between Trained and Naïve Horses

There was no significant difference between Naïve and Trained Control HGS (T-Test: T = 0.347, N1 = 13, N2 = 12, P = 0.733). There was also no difference in HGS between Trained and Naïve horses when wearing the Dually (T-Test: T = 1.42; N1 = 12, N2 = 14, P = 0.179). Further, there was no difference in HGS between Dually and Control in Naïve horses, when considering re-scored images (Mann Whitney: V = 13, N = 8, P = 0.528). When wearing the Dually™, Trained horses did not have significantly lower IRT changes, compared to Naïve horses (T-Test: T = 0.448, N1 = 14, N2 = 16, P = 0.251). When wearing the Dually™, Trained horses did not cross the obstacle significantly more quickly that Naïve horses (Mann Whitney: U = 188, N1 = 19, N2 = 16, P = 0.239). Trained horses did show significantly more proactive behaviour than Naïve horses when wearing the Dually™ (T-Test: T = -3.904, N1 = 13, N2 = 9, P = 0.002) and a strong effect was observed (D = 0.753). No difference in proactivity was observed between Trained and Naïve horses in the Control (Mann Whitney: U = 77, N1 = 14, N2 = 11, P = 1). No other variables differed between Trained and Naïve horses (Table 5).

Table 5. Differences in behaviour and physiology between Trained and Naïve horses for Dually and Control. Independent T-Tests (TT) and Mann Whitney U-Tests (MW) were conducted as appropriate for normality.

Variable	Treatment	Mean/ Median	SD/IRQ	Test	U/T	P	Effect Size	N
HGS	Naïve Control	1.9	±1.9	TT	0.347	0.733	0.082	13
	Trained Control	1.7	±0.93					12
	Naïve Dually	2.96	±2.27	TT	1.42	0.179	0.366	12
	Trained Dually	1.99	0.75					14
IRT Change	Naïve Control	-0.44	±1.05	TT	1.181	0.251	0.439	14
	Trained Control	-0.06	±0.7					16
	Naïve Dually	-0.2	±0.81	TT	0.448	0.658	0.163	14
	Trained Dually	-0.06	±0.82					16
Crossing Time	Naïve Control	31	132.5	W	174	0.474	-0.119	19
	Trained Control	20.7	47.75					16
	Naïve Dually	40	128.5	W	188	0.239	-0.196	19
	Trained Dually	23.3	57.5					16
% Pro-activity	Naïve Control	17.15	15.32	W	77	1	0	14
	Trained Control	10.72	63.7					11
	Naïve Dually	15.65	±14.91	TT	-3.904	**0.002**	**0.753**	13
	Trained Dually	53.3	±26.12					9

4. Discussion

The aim of the present study was to investigate how trained horses to respond to the pressure of the Dually™ headcollar and how this affects compliance and stress in a novel handling test. The impact of the Dually™ on stress physiology during training and testing was also assessed. Following training, horses were asked to complete two novel handling tests, once with the line attached to the side-ring and once with the line attached to the standard under chin ring as a control. Results suggest the Dually™ may have a negative effect on compliance but does not cause welfare concerns in horses trained to respond to the pressure/release mechanism.

During the novel test, Trained horses in the Dually™ were not significantly quicker to cross the novel object than horses in the Control headcollar setting. Further, Trained horses did not cross more quickly than Naïve horses. The first Dually™ study also demonstrated no difference in crossing time between horses wearing the Dually™ and those wearing a control headcollar (Ijichi et al., 2018). One of the limitations to the first study was that subjects had no prior training in the Dually™, therefore it could be expected that training would improve compliance. It is generally agreed that training horses to respond to handler signals via stimulus generated by pressure from a headcollar is an effective way to achieve compliance (McLean, 2005). However, there was a tendency for Trained horses to be more proactive in the Dually™ than the Control and significantly more so than Naïve horses in the Dually™. No difference was seen for proactivity between Trained and Naïve horses for the Control setting, indicating that differences seen in the Dually cannot be explained by the different sample of horses. This suggests that training in fact increased resistance to the device, rather than improving it as the horse learns how to release the pressure. Taken together, this indicates that the Dually™ does not improve compliance during handling. It is not clear whether further training would extinguish or exacerbate this proactive response.

It may be that three training sessions were not sufficient to significantly alter the effect of the Dually™. However, subjects experienced an average of 157 (±22) attempts in this time and during training all horses in the study were compliant and able to consistently offer the desired response. Another possibility is that the three-minute handling challenge was not long enough for the effect of the Dually™ to be observed. This is contradicted by the fact that all but one horse crossed within this time. A counter explanation for the lack of effect of the Dually™ is that the handling tests were not aversive enough. However, most horses (60%) resisted crossing the obstacle in the current study. Further, LF/HF was elevated, whilst RMSSD decreased, indicating that the handling tests were inducing observable arousal. More aversive tests may not be considered ethically appropriate within the context of research. Finally, proponents of the device might explain this lack of improvement following training by noting that we did not perform "join-up" during training. However, multiple sources of evidence indicate this is not a useful training approach for building bond (Henshall et al., 2012) and does not generalise to other contexts (Krueger, 2007).

In the previous research, HGS scores were significantly higher in the Dually™ compared to the control (Ijichi et al., 2018). However, the scorer was not blind to treatment, as these cannot easily be obscured from the photos without limiting how clearly the face can be observed. In the current experiment, a hypothesis-blind rater was used to resolve this limitation. In the current study, there was no difference in HGS between Dually™ and Control in Trained horses. Whilst this might suggest that training reduces the discomfort caused by the Dually, there was no difference in HGS between Trained and Naïve subjects during Dually use. This indicates that it is not training per se that explains this finding. In fact, reanalysed HGS for Naïve horses did not show a significant difference between Dually and Control, challenging the finding of the original paper. This is likely to be the result of including all images (rather than a random sample) and calculating HGS by averaging each Facial Action Unit (FAU) and then totalling these (rather than using percentage to account for missing FAU). Whilst HGS were still higher for Dually compared to Control this was no longer significant. Further research could be conducted to observe behaviour and HGS longitudinally in horses being tested in the Dually for the first time compared to after a period of training.

Although the Dually™ had a potentially negative effect on compliance, there was no effect of training on stress indicators. There was no difference in IRT, RMSSD, LF/HF or salivary cortisol between Dually™ and Control, suggesting the Dually™ does not reduce welfare within a 3-minute handling challenge when compared to a standard headcollar. This does not contradict findings that the Dually caused greater proactivity, as proactive behaviour does not necessarily indicate higher arousal (Munsters et al., 2013; Squibb et al., 2018; Yarnell et al., 2013). Similar stress profiles between Dually and Control supports the observation in the original research which indicated there was no difference in IRT between Dually™ and Control in Naïve horses, despite higher HGS scores (Ijichi et al., 2018).The current study measured cortisol in addition to the measures used by Ijichi et al (2018) but it is possible that peak cortisol changes would have been captured sooner than the 10 minute latency used here (Contreras-Aguilar et al., 2019). However, no other stress indicator changed as a result of testing and IRT did not differ between Trained and Naïve horses. However, it is worth considering that these indicators of arousal might alter if the testing lasted longer than 3 minutes. For example, studies investigating the effects of tight noseband, which apply pressure to the same anatomical structures, observed horses for 10 minutes (Fenner et al., 2016; McGreevy et al., 2012). It is important to know whether longer handling sessions more representative of typical behaviour modification sessions do result in stress. Indeed, average RMSSD significantly decreased whilst LF/HF significantly increased during Training compared to a stabled baseline. These HRV variables suggest that training in the Dually™ headcollar caused observable arousal (Stucke et al., 2015), though this was not seen in IRT or salivary cortisol changes. This might be explained by the fact that baseline measures for HRV were taken in the stable but IRT and salivary cortisol were taken within the research arena. This may have caused a feed-forward anticipatory stress response to raise the baseline values. However, this was important to account for diurnal fluctuations in cortisol (Hoffis et al., 1970). Further, it is not clear whether the Dually™ caused more arousal than the same training in a standard headcollar, as Control training sessions were not conducted.

5. Conclusion

The findings of the current study suggest that the Dually™ does not improve compliance in trained horses as horses do not cross more quickly compared to a standard headcollar. In fact, potentially dangerous proactive behaviour was increased in the Dually™ and is exacerbated by training, rather than diminishing this response. It should be noted that the device does not appear to cause more stress or discomfort than standard headcollars in Trained horses, though the short testing time may not be sufficient to detect an effect of the headcollar on arousal. Therefore, while the efficacy of the device is questionable, it does not appear to cause poorer welfare and if owners perceive that it gives them more control this may justify its use.

Acknowledgements

We are indebted to Anna Gregory, Cath Hake and Jake Bromley-Fowles for facilitating this research on the Brackenhurst yard. Salivary cortisol analysis was supported by the European Regional Development Fund.

Author Contributions

The idea for this paper was conceived by Carrie Ijichi; the experiment was designed by Carrie Ijichi and Hayley Wild; data was collected by Carrie Ijichi, Hayley Wild, Heather Cameron-Whytock, Samuel White, Aurelie Jolivald, Sarah Hallam and Lauren Birkbeck; analysis was done by Carrie Ijichi, Francesca Dai; Emanuela Dalla Costa, Hayley Wild, Alex Bordin, Gareth Starbuck and Kelly Yarnell; statistical analysis was done by Carrie Ijichi; the paper was written by Carrie Ijichi and drafted by all authors. The authors of this manuscript have no conflict of interest to declare and no funding bodies to acknowledge.

References

Brubaker, L., Udell, M.A.R., 2016. Cognition and learning in horses (Equus caballus): What we know and why we should ask more. Behav. Processes 126, 121–131. https://doi.org/10.1016/j.beproc.2016.03.017

Church, J.S., Hegadoren, P.R., Paetkau, M.J., Miller, C.C., Regev-Shoshani, G., Schaefer, A.L., Schwartzkopf-Genswein, K.S., 2014. Influence of environmental factors on infrared eye temperature measurements in cattle. Res. Vet. Sci. 96, 220–226. https://doi.org/10.1016/j.rvsc.2013.11.006

Dalla Costa, E., Minero, M., Lebelt, D., Stucke, D., 2014. Development of the Horse Grimace Scale (HGS) as a pain assessment tool in horses undergoing routine castration. PLoS One 9, e92281.

Doherty, O., Conway, T., Conway, R., Murray, G., Casey, V., 2017. An objective measure of noseband tightness and its measurement using a novel digital tightness gauge. PLoS One 12, e0168996. https://doi.org/10.1371/journal.pone.0168996

Fenner, K., Yoon, S., White, P., Starling, M., McGreevy, P., 2016. The Effect of Noseband Tightening on Horses' Behavior, Eye Temperature, and Cardiac Responses. PLoS One 11, e0154179. https://doi.org/10.1371/journal.pone.0154179

Field, A., Miles, J., Field, Z., 2012. Discovering Statistics Using R. SAGE Publications Ltd, London.

Hoffis, G., Murdick, P., Tharp, V., Ault, K., 1970. Plasma concentrations of cortisol and corticosterone in the normal horse. Am. J. Vet. Res. 31, 0179–1387.

Hughes, T., Creighton, E., Coleman, R., 2010. Salivary and fecal cortisol as measures of stress in horses. J. Vet. Behav. Clin. Appl. Res. 5, 59–60.

Ijichi, C., Collins, L.M., Creighton, E., Elwood, R.W., 2013. Harnessing the power of personality assessment: Subjective assessment predicts behaviour in horses. Behav. Processes 96, 47–52. https://doi.org/10.1016/j.beproc.2013.02.017

Ijichi, C., Green, S., Squibb, K., Carroll, A., Bannister, I., 2019. Zylkéne to Load? The effects of alpha-casozepine on compliance and coping in horses during loading. J. Vet. Behav. 30, 80–87. https://doi.org/10.1016/j.jveb.2018.12.009

Ijichi, C., Tunstall, S., Putt, E., Squibb, K., 2018. Dually Noted: The effects of a pressure headcollar on compliance, discomfort and stress in horses during handling. Appl. Anim. Behav. Sci. 205, 68–73.

Ille, N., Erber, R., Aurich, C., Aurich, J., 2014. Comparison of heart rate and heart rate variability obtained by heart rate monitors and simultaneously recorded electrocardiogram signals in nonexercising horses. J. Vet. Behav. Clin. Appl. Res. 9, 341–346. https://doi.org/10.1016/j.jveb.2014.07.006

McGreevy, P., McLean, A., 2007. Roles of learning theory and ethology in equitation. J. Vet. Behav. Clin. Appl. Res. 2, 108–118.

McGreevy, P., Warren-Smith, A., Guisard, Y., 2012. The effect of double bridles and jaw-clamping crank nosebands on temperature of eyes and facial skin of horses. J. Vet. Behav. Clin. Appl. Res. 7, 142–148. https://doi.org/10.1016/j.jveb.2011.08.001

McLean, A.N., 2005. The positive aspects of correct negative reinforcement. Anthrozoos A Multidiscip. J. Interact. People Anim. 18, 245–254. https://doi.org/10.2752/089279305785594072

Munsters, C., Visser, K., van den Broek, J., Sloet van Oldruitenborgh-Oosterbaan, M.M., 2013. Quantifying stress in experienced and inexperienced mounted police horses, using heart rate, heart rate variability, behavior score and suitability score. J. Vet. Behav. Clin. Appl. Res. 8, e16–e17. https://doi.org/10.1016/j.jveb.2012.12.037

RStudio Team (2020). RStudio: Integrated Development for R. RStudio, PBC, Boston, MA URL http://www.rstudio.com/.

Roberts, M., 1999. Controlling halter for animals. Google Patent No. US6062005

Squibb, K., Griffin, K., Favier, R., Ijichi, C., 2018. Poker Face: Discrepancies in behaviour and affective states in horses during stressful handling procedures. Appl. Anim. Behav. Sci. 202, 34–38. https://doi.org/10.1016/j.applanim.2018.02.003

Stucke, D., Große Ruse, M., Lebelt, D., 2015. Measuring heart rate variability in horses to investigate the autonomic nervous system activity - Pros and cons of different methods. Appl. Anim. Behav. Sci. 166, 1–10. https://doi.org/10.1016/j.applanim.2015.02.007

von Borell, E., Langbein, J., Després, G., Hansen, S., Leterrier, C., Marchant-Forde, J., Marchant-Forde, R., Minero, M., Mohr, E., Prunier, A., Valance, D., Veissier, I., 2007. Heart rate variability as a measure of autonomic regulation of cardiac activity for assessing stress and welfare in farm animals - A review. Physiol. Behav. 92, 293–316. https://doi.org/10.1016/j.physbeh.2007.01.007

Williams, F., Ashby, K., 1995. Horse Related Injuries, Hazard, Monash University Accident Report Centre.

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. https://doi.org/10.1016/j.physbeh.2013.05.013