Associations Between Commonly Used Apparatus and Conflict Behaviors Reported in the Ridden Horse in Australia

Condon, V. M.; McGreevy, P. D.; McLean, A. N.; Williams, J. M.; Randle, Hayley

Published in:
Journal of Veterinary Behavior: Clinical Applications and Research

Publication date:
2021

ASSOCIATIONS BETWEEN COMMONLY USED APPARATUS AND CONFLICT BEHAVIORS REPORTED IN THE RIDDEN HORSE IN AUSTRALIA

Condon, V.M.[1*], McGreevy, P.D.[2], McLean, A.N.[3], Williams, J.M.[4], Randle, H.[1]

[1]School of Animal and Veterinary Science, Charles Sturt University, Wagga Wagga, NSW, 2678.

[2]Sydney School of Veterinary Science, University of Sydney, NSW 2006, Australia

[3] Equitation Science International, 3, Wonderland Ave, Tuerong, VIC 3915, Australia

[4]Hartpury University, Gloucester, GL19 3BE, United Kingdom.

[*]Corresponding author: victoriacondon@bigpond.com

Abstract

Equestrian equipment is often used to maintain control of horses while riding, training or handling them and therefore to optimize human and horse safety. However, equipment that has been incorrectly selected or inappropriately used may result in horses exhibiting conflict-related behavior. Characterising associations between apparatus use and unwelcome horse behavior could benefit both horse welfare and human safety by elucidating the ontogeny of undesirable equine responses and promoting the ethical use of equipment. The current study explored associations between commonly used apparatus and the reported behaviors of the ridden horse, using an online survey that attracted 1101 Australian respondents. Chi-square tests of association revealed thirteen (9.1%) significant relationships between any unwanted behavior and single items of apparatus used during riding. Analysis of combinations of apparatus that impose aversive stimuli (e.g. harsh bits for deceleration, and spurs for acceleration) revealed that 37.19% of combinations of such items were significantly related to unwanted/conflict equine behaviors. Additional analysis demonstrated no significant difference in the unwanted behavior of horses ridden in bitless versus bitted bridles. This study has demonstrated associations between unwanted ridden behaviors and type of apparatus used, particularly when multiple items are used to apply opposing aversive stimuli. Whilst these results do not imply causation, they may be used to better inform equestrians' choice and ethical use of apparatus.

Keywords: equine; welfare; tack; gear; bit; whip

Introduction

The role of the domestic horse has evolved with the advance of civilisation. Once used for locomotion and agricultural power, horses now more commonly play roles in sport, leisure and companionship (McGreevy et al., 2018b). The empirical study of horse handling, training and riding is central to the nascent discipline of equitation science (McGreevy, 2007) and, as the impact of human interventions on domestic horse welfare is being scrutinized (McGreevy et al., 2018a), the exploration of what constitutes ethical equitation (Jones and McGreevy, 2010; Randle and Waran, 2019) is increasingly important.

Horse-riding involves the rider establishing stimulus control over the horse's locomotory responses (McLean and McLean, 2008) commonly using equipment, such as bits, bridles, whips and spurs, to train the horse to respond in a particular way via negative reinforcement (NR). In NR, the horse learns locomotory responses to these pressure cues (historically called aids) because they are associated with the cessation of aversive stimuli. In correct NR, pressure cues should rapidly shrink to light signals (the discriminative stimuli), and light cues should precede any stronger pressure (McGreevy et al., 2018a). Whilst horses are skilled at responding to NR (McLean and Christensen, 2017), human error in the application and removal of pressure, such as poor and/or inconsistent timing, can lead to training-related stress, confusion, poor training outcomes, and compromised welfare. Simultaneous application of acceleration and deceleration pressures of similar intensity increases the likelihood of overshadowing and, ultimately, habituation to the least salient pressure (McGreevy et al., 2018b). The simultaneous application of contradictory cues, e.g. that trigger both acceleration and deceleration is contraindicated (McLean and McLean, 2008).

To cope with stressors, animals may attempt to remove themselves from discomfort, confront the stressors, or adapt to them (Moberg and Mench, 2000). If unable to resolve conflicting or unrelenting stimuli, horses may exhibit behavior indicative of a state of conflict or ultimately habituate to the stimuli (McLean and McGreevy, 2010b). Mills and Marchant-Forde (2010) describe conflict behaviour as “a category of stress-induced behavior changes that arise from conflicting motivations”. In the ridden horse these may manifest as body tension, rearing, bolting, bucking or shying. Both conflict-related behaviors and inadvertent habituation are undesirable in the context of the human-horse relationship. In many competitive equestrian disciplines, such responses are associated with poor performance and many directly attract penalties.

Doherty et al. (2017c) suggested some indicators of a horse’s pain, stress or fear go unnoticed by the rider. Behaviors such as shying may also be misinterpreted as disobedience or a merely reflection of the horse’s temperament, particularly in the context of equitation. Known conflict related behaviors, such as rearing and bucking, may be clearly recognisable, but interpretation of both these behaviors and their etiology is complex because their presentation may be influenced by numerous factors, including rider/handler perception (Williams and Tabor, 2017). Behavioral responses to stimuli reflect inter-individual differences in reactivity and sensitivity that are influenced by both genetic (Le Scolan et al., 1997; Hausberger and Muller, 2002) and environmental factors or context (Søndergaard and Ladewig, 2004; Hausberger et al., 2004).

Clinical indicators of relatively subtle pain responses, such as teeth-grinding, reluctance to perform required movements and/or gait changes, muscular tension or atypical body posture may go undiagnosed and minor problems may potentially escalate into more serious welfare issues (Hall et al., 2013; Randle and Waran, 2017). Many common items of apparatus used in equitation, such as bits and nosebands, are designed to apply aversive pressure to control and direct horses’ movement. Other equipment, such as martingales or auxiliary reins, can use mechanical principles, such as leverage, to amplify or redirect the pressure already being applied with the intention of preventing the horse from evading the original pressures (McGreevy et al., 2018b). Even the most seemingly innocuous equipment that are commonly considered ‘basic’ and even ‘kind’, such as the snaffle bit, can be a source of abuse if used to apply pressure inappropriately, excessively or inconsistently. In equitation, the human is responsible for the force and duration of aversive pressure applied. However, studies of the direct impact of apparatus-related pressure upon horse behavior are in their infancy (Doherty et al., 2017a; Randle et al., 2017). That said, numerous behaviors exhibited by the horse under saddle are seldom observed outside the ridden context and should be valued among other potential animal-based indicators of poor welfare (Randle and Waran, 2017). Increased interest in animal welfare has drawn attention to potentially questionable ethics around some aspects of horse use in sport and recreation (McLean and McGreevy, 2010a, McLean and McGreevy, 2010b) and public scrutiny is intensifying with regard to social license to operate within racing (McGreevy and McManus, 2017) and many other sectors of the equine industry. Within this context, many training methods and items of equipment, some of which have justifications based chiefly on tradition, are now being questioned more deeply than ever.

Equestrian apparatus may, by design, be intrinsically aversive particularly when making physical contact with extremely sensitive areas of the horse. Possibly exacerbated by selective breeding, individual horses show considerable variation in sensitive loci. For example, some horses are extremely ‘head-shy’ or ‘girth-shy’ while other individuals are seemingly oblivious to tactile contact in these regions (McGreevy et al., 2018b). Most of the equipment used on horses has been designed to benefit humans, providing comfort, stability and a means of controlling the ridden horse. All items designed to transmit pressure in equitation can be sources of abuse if deployed inappropriately, excessively or inconsistently.

Paleopathological studies have identified effects of horse-riding apparatus in ancient skeletal remains of horses, including bit-wear on the jaw and teeth (Anthony and Brown, 2015) and abnormalities of the postcranial skeleton indicative of the effects of riding and traction (Levine et al., 2000). There is growing evidence that contemporary equipment used to control horses can inadvertently harm them (Mata et al., 2015; Crago et al., 2019; Uldahl and Clayton, 2019; Weller et al., 2020).

In a survey of equestrian apparatus use in Australia, Hill et al. (2015) noted the use of milder bits, designed to be less aversive than average, was associated with a reduction in the use of devices used to accelerate the horse, notably whips and spurs. Correspondingly, relatively severe bits such as curbs, designed specifically to apply stronger deceleration pressures, were often used in conjunction with these accelerator devices. These findings raise questions about the use of such apparati with opposing effects, notably those designed to amplify pressures for either acceleration or deceleration, being used simultaneously. When operator error, such as the delayed release of pressure, habituates horses to pressure cues, then escalating the aversiveness with more severe items of gear offers only a short-term solution at best. Unless the riders' errors are remediated, the perceived need for yet more severe gear will eventually emerge.

Hockenhull and Creighton (2013) investigated the use of items of apparatus, with a focus on so-called artificial aids, such as martingales, spurs and whips, in relation to owner-reported prevalence of ridden behavior problems in a sample of UK leisure horses (n=1,326). Ninety-one percent of the horses were reported as displaying unwanted behavior when ridden in the week preceding data submission. Furthermore, 78% of respondents reported using one or more artificial aids when riding or handling the horse on the ground (Hockenhull and Creighton, 2013). The widespread reported use of these artificial aids is of concern if they escalate force without requiring optimal operation that avoids excessive force and poor timing. While not inferring a direct causal relationship between apparatus and behavior, Hockenhull and Creighton (2013) proposed that behavioral conflict may be exacerbated by owners' attempts to address behavioral issues with apparatus that increases the intensity of controlling signals.

The complex etiology of conflict behavior must be considered in the context of human-horse relationships (McGreevy et al., 2009), most notably those associated with training and performance. In addition, the collection of valid, reliable and scientifically robust data is difficult, being complicated by numerous dynamic, confounding variables that exist within the horse-human dyad. It is worth considering that unwanted behaviors may cluster together suggesting inter-relatedness. It is therefore plausible that they may be triggered by, or result from, the use of similar items of apparatus. Indeed, it is also possible that clusters of locomotory behaviors may trigger corresponding clusters of reactions from riders that may include the use of items of gear that amplify pressure. Researchers and practitioners alike are now questioning whether, in the absence of owners' having a thorough understanding of the consequences of using increasingly aversive stimuli, behavioral conflict may be exacerbated by use of such equipment (Hall et al., 2013; Heleski et al., 2009). Statistical testing can be used to identify direct associations between apparatus use and horse behavior. If significant associations between items of apparatus and undesirable horse behaviors exist, identification of these relationships can benefit both horse welfare and human safety by informing owners about the range of effects and consequences items of apparatus can have and providing information about its ethical use.

The effect of bit pressure on the horse is complex, with multiple physical outcomes from increased rein tension, beyond pressure on the tongue and diastema (Doherty et al., 2017b). Despite researchers such as Cross et al., (2017) positing the physical effect of bit operation can be inferred from laboratory testing, multiple variables may potentially affect the physical impact on the horse, including the bit itself, the rider's hands and the anatomy of the horse's mouth. Meanwhile, arguments for bitless bridles appear frequently in industry and practitioners' media (Cook and Mills, 2009), with the suggestion that they are inherently more ethical than their bitted counterparts (Mellor and Beausoleil, 2017). Hence, the relationships among bitted bridles and unwanted horse behaviors may be of particular interest. The current study investigated possible relationships among conflict behaviors reported in the ridden horse in Australia and associations with commonly used apparati.

Materials and methods

The raw data were collected by Hill et al. (2015) via an online questionnaire designed to audit apparatus used to apply aversive stimuli to ridden horses and ponies in Australia. The study was approved by the University of Sydney Research Ethics Committee (approval number 12396); the questionnaire was available between 20th July 2012 and 26th September 2012. Further details and preliminary results of this study have previously been published (Hill et al., 2015). A copy of the survey is available as supplementary material.

The questionnaire asked for details of riding style and equipment used. A further set of data about horse behaviors that were exhibited when ridden were collected using the same instrument. Since the relationships between apparatus used and occurrence of unwanted/undesirable behaviors were not analysed by Hill et al. (2015), the current study analysed data relating to ridden behavior and accompanying apparatus to identify associations among them.

Horse ridden behavior

Respondents were asked about their horses' ridden behavior and provided with five possible options for each response. The 11 ridden behavior statements provided were: *this horse pulls its head forward and down when you are riding, this horse gets excited when he anticipates cantering, this horse never bolts, this horse never bucks, this horse never rears, this horse never shies, this horse is very sensitive to my leg cues, this horse needs strong leg cues to get into canter, this horse is very easy to steer, this horse stops very easily,* and *this horse loves to jump obstacles.* Respondents were asked to use a 5-point Likert response scale (from 0-4, where 0=never, 4=always) to indicate the frequency with which the horse expressed the 11 listed behaviors. A "not applicable" option was also provided to reduce the possibility of a non-response. To facilitate analysis, low, nominal values were merged and these data recoded to binary responses (rankings 0 and 1 became 'no", and rankings 2, 3 and 4 became 'yes'). Responses left blank were excluded from the analysis.

Apparatus applied

Apparatus-related data included the broad range of equipment commonly used during equitation activities. Respondents were asked about 13 categories of common apparatus and the type of apparatus used within each category was recorded. (The number of options are shown in parentheses.) Apparatus categories and type were: bit style (n=4); bit thickness (n=5); bit material (preliminary analysis identified n=57 variations of bit material composition); bit mouthpiece type (n=15); shank length (n=7); number of reins (n=3); chain or lip strap (n=6); noseband type (n=35); auxiliary reins (n=5); martingale type (n=5); other auxiliary equipment (n=7); spur type (n=9) and whip type (n=7). In any instance where the respondent gave more than one response, the item of apparatus considered most aversive/restrictive according to criteria defined by McGreevy et al. (2018a) was recorded. Responses that did not provide information on the type of apparatus used in each category were eliminated from the analysis. Bridle use was also recorded as either bitted (n=729 respondents) or bitless (n=132 respondents).

Statistical analysis

Using data shown in Figure 1 and Table 2, categories of responses detailing behavior (n=11) and apparatus (n=13) were tabulated into 143 2 × 2 contingency tables using Minitab Statistical Software version 18 (Minitab, LLC, State College, Pennsylvania, USA). Chi-square tests of association were conducted on all variables. Initial analysis revealed no associations between any behavior and single item of apparatus. Data on types of apparatus were then combined to reduce the number of variables in each category, whilst still meeting validity requirements for chi-square tests. These then became the following 13 factors: bit style (n=3: bitless, snaffle, double/Pelham); bit thickness (n=3: thin, medium, thick); bit material (n=2: metal, other); bit mouthpiece type (n=3: jointed, no joint, other); shank length (n=2: shank, no shank); number of reins (n=2: one pair, more than one pair); chain or lip strap (n=2: either chain or strap, none); noseband type (n=3: halter, cavesson, other); auxiliary reins (n=2: draw reins, other); martingale type (n=2: standing martingale, other); other auxiliary equipment (n=2: martingale only, other); spur type (n=2: smooth, other) and whip type (n=2: dressage whip, other). A further series of chi-square tests were run on these modified 143 contingency tables. To investigate the possibility of an association between bitted *versus* bitless bridle use and behavior, chi-square tests were conducted using a total of 861 responses; the apparent shortfall being attributed to 237 participants (21.6%) not providing data in response to the question of bitted/bitless bridle use.

The next analysis explored any associations between the expression of ridden behavior and apparatus applying opposing aversive stimuli when used in combination. Apparatus categories were concatenated to include accelerator devices (n=2), specifically whips and spurs, used concurrently with decelerator devices (n=11), such as bits, nosebands, martingales and auxiliary reins (draw reins, side reins, Whitaker reins, Pessoa® Training System, Market Harborough, tie-down, Chambon).

Contingency tables (n=121) were created relating behaviors (n=11) to apparatus that allows riders to apply opposing aversive stimuli (n=11) and chi-square tests were performed on these behavior and (concatenated) apparatus data. Significance was set as $P < 0.05$ *a priori*.

Binary logistic regression

A series of binary logistic regression models were constructed to identify risk factors associated with the reporting of specific undesirable behaviors during ridden work. The variables to be fitted in the multivariable models were first assessed using univariate analysis of binary response variables: *pulls-head-forward-down:* yes or no, *very-sensitive-to-leg-cues*: yes or no, *needs-strong-canter-cues*: yes or no, *gets-excited-anticipates-canter*: yes or no, *shies*: yes or no, *easy-to-steer*: yes or no, *easy-to-stop*: yes or no, and *loves-to-jump-obstacles:* yes or no. A variable with an alpha value of < 0.10 was considered eligible for use in building the multivariable models (Bailey et al., 1997). No significant factors were related to the behaviors *pulls-head-forward-down*, *shies* or *easy-to-steer*, so five multivariate models were built (Model A: *Very-sensitive-to-leg-cues*; Model B: *Needs-strong-canter-cues*; Model C: *Gets-excited-anticipates-canter*; Model D: *Easy-to-stop*; Model E: *Loves-to-jump-obstacles;* see Table 1). Each model was fitted using a backward stepwise process that excluded variables with a likelihood ratio test significance of $P < 0.05$ (Parkin et al., 2006; Smith et al., 2018). For each step in the multivariate model building process, the effect of removal of variables was assessed using a likelihood ratio chi-square test of model coefficients ($P < 0.05$) to check that the new model was an improvement on the baseline model (Williams et al., 2013; Smith et al., 2018). This ensured that variables with a significant impact on the model were not excluded from further analysis. Hosmer-Lemeshow goodness-of-fit tests ($P > 0.05$) were also used to evaluate the fit of the model (Williams et al., 2013; Smith et al., 2018). The predictive ability of the final model was investigated using receiver operating characteristic (ROC) curve analysis (Williams et al., 2013; Smith et al., 2018). The risk of the behavior in question for each model being reported at an increased prevalence was compared using the odds ratio (OR) and associated 95% confidence intervals (CI).

Table 1: Multivariate models assessing the presence of specific undesirable behaviors during ridden work. Five multivariate models were built (Model A: *Very-sensitive-to-leg-cues*; Model B: *Needs-strong-canter-cues*; Model C: *Gets-excited-anticipates-canter*; Model D: *Easy-to-stop*; Model E: *Loves-to-jump-obstacles*)

Model	**Factors taken forward to the final model (*P* < 0.1)**
Model A: Very-sensitive-to-leg- cues	Noseband type - Headcollar/halter (1- reference), Plain cavesson (2), Other (3) Noseband type and whip - Headcollar/halter (1- reference), Plain cavesson (2), Other (3), Whip (4), Headcollar/halterWhip (5), Plain cavessonWhip (6), OtherWhip (7)
Model B: Needs-strong-canter-cues	Lipstrap or chain or None - Yes chain or strap (1- reference), Neither chain or lip strap(2), Spur type - Smooth spurs(1- reference), Other(2) Other Auxiliary equipment and spurs – spurs (1 - reference), Martingale only(2), Other(3)
Model C: Gets-excited-anticipates- canter	Shank length and spurs – shank (1- reference), No shank(2), spurs(3), shankspurs(4), No shankspurs(5), Martingale type and spurs - spurs(1- reference), Standing martingale(2), Other martingale(3), Standing martingalespurs (4), Other martingalespurs (5) Other auxiliary equipment and spurs - spurs(1- reference), Martingale only(2), Other(3) Bit thickness and spurs – Any thickness bit no spurs (1- reference), Thin + spurs(2), Medium+spurs(3), Thick+spurs(4)
Model D: Easy-to-stop	Bit style and spurs - Any bit no spurs(1- reference), Snaffle+spurs(2), Double/Pelham +spurs(3), Number reins used and spurs - Only one(1 - reference), One setspurs(2), one or morespurs(3)
Model E: Loves-to jump-obstacles	Noseband type and spur Bit style and whip - snaffle (1 – reference), double/pelham (2), Bitless (3), snaffle + whip (4), double/pelham + whip (5), bitless+whip (6), whip (7)

	Chain/lip strap and whip – neither chain, lip strap or whip (1-reference), neither chain or lip strap (2), whip (3), chain or strap (4), chain strap or whip (5) Aux reins and whip - draw reins (1-reference), draw reins and whip(2), whip (3), other whip (4), other (5) Martingale type and whip - standing martingale (1 – reference), standing martingale + whip (2), whip (3), other martingale (4), other martingale + whip (5)

Results

Data used for this study were derived from a total of 1101 survey responses by respondents over 18 years of age: 82.6% females (n=910), 17.1% males (n=188) and 0.3% (n=3) who did not disclose gender. The questionnaire collected details of main riding discipline and equipment/apparatus used and included a set of 11 behaviors horses exhibit when ridden.

Horse ridden behavior

Of the 1101 respondents, the majority (n=997) reported at least one problematic ridden behavior. Responses from the remaining 4 respondents were not used in further analysis. Horse ridden behavior frequency adjusted to binary responses are shown in Figure 1.

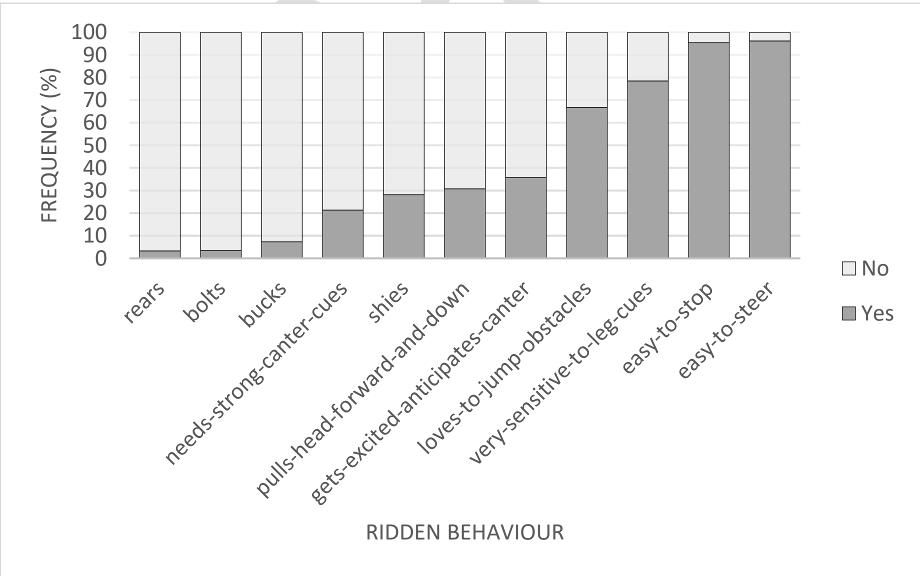


Figure 1. Frequency of horse ridden behaviors reported by 1101 respondents

Apparatus applied

The number of apparatus variables for both preliminary and secondary analyses are shown in Table 2, along with number of responses to each category of apparatus. Table 2 also shows the percentage of non-responses recorded as “no data”.

Table 2. Types of apparatus reported by 1101 respondents. Preliminary analyses were conducted on all variables using raw data from Hill et al. (2015). Data on types of apparatus were subsequently combined to reduce the number of levels and secondary analysis conducted. For example, in the preliminary analysis, variable bit style included 4 levels (bitless; snaffle; Weymouth/double bridle and curb/shank (Pelham)) and in the secondary analysis due to combining categories, variable Bit style had only 3 variables (bitless, snaffle and double bridle/Pelham).

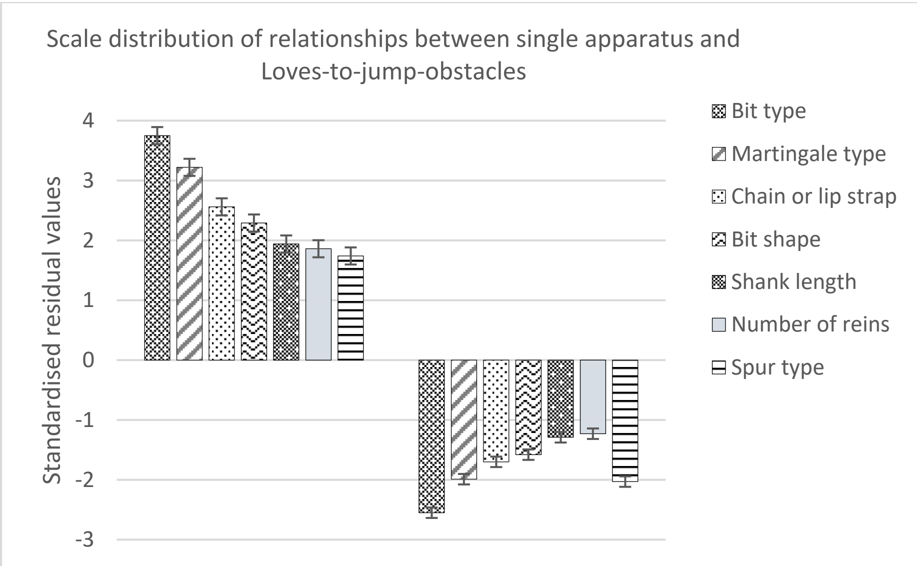
Apparatus = VARIABLE	Number of levels in preliminary analysis	Number of levels in secondary analysis	Responses	Responses (%)	*(no data)	*(%) of no data	Total responses
Bit style	4	3	861	78.42	237	21.58	1098
Bit thickness	5	3	737	68.24	343	31.76	1080
Bit material	57	2	740	68.52	340	31.48	1080
Bit mouthpiece type	15	3	742	68.70	338	31.30	1080
Shank length	7	2	581	53.80	499	46.20	1080
Number reins used	3	2	614	56.85	466	43.15	1080
Chain or lip strap	6	2	599	55.46	481	44.54	1080
Noseband type	35	3	652	59.38	446	40.62	1098
Auxiliary reins	5	2	78	7.09	1022	92.91	1100
Martingale type	5	2	197	17.91	903	82.09	1100
Other auxiliary equipment	7	2	74	6.73	1026	93.27	1100
Spur type	9	2	355	32.27	745	67.73	1100
Whip type	7	2	499	46.12	583	53.88	1082

Statistical data analysis – behaviors associated with a single item of apparatus

Preliminary analysis revealed no associations between any behavior and any single item of apparatus. Secondary analysis (i.e. with combined response categories) revealed 13 (9.1%) significant associations between an identified behavior and a single item of apparatus used during riding ($P<0.05$), as shown in Figure 2.

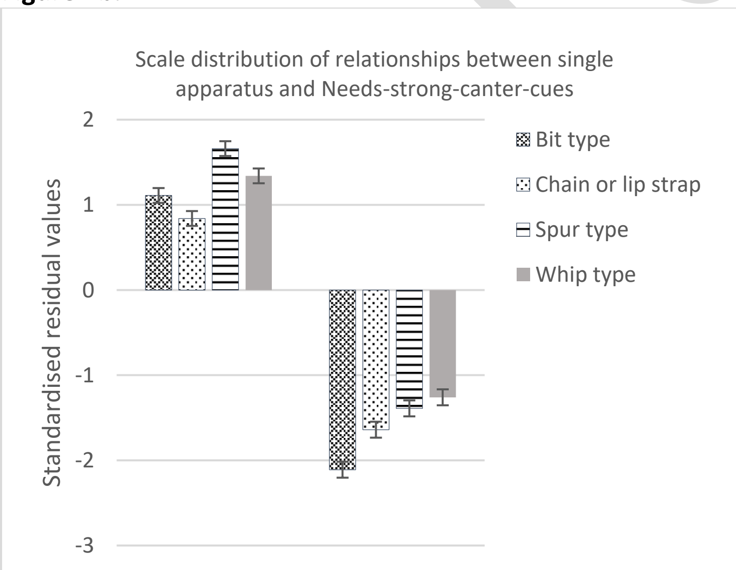
Of the 11 surveyed behaviors, only 4 (36.36%): *pull-head-forward-and-down; very-sensitive-to-leg-cues; needs-strong-canter-cues* and *loves-to-jump-obstacles*, showed significant associations with the use of a single item of apparatus.

Figure 2a.



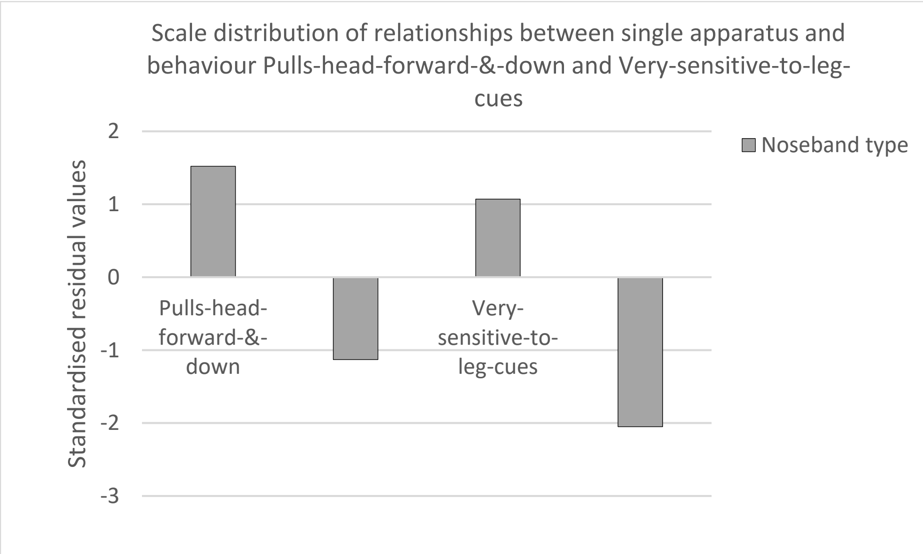
Example, *Love-to-jump-obstacles*/chain or lip strap: more horses do not *Love-to-jump-obstacles* with a chain or lip strap than expected (SR = 2.56), fewer horses *Love-to-jump-obstacles* with a chain or lip strap than expected (SR = -1.70)

Figure 2b.



Example, *Need-strong-canter-cues*/whip type: more horses *Need-strong-canter-cues* with other type whip than expected (SR = 1.34), fewer horses *Need-strong-canter-cues* with a dressage whip than expected (SR = -1.27)

Figure 2c.



Example, *Pulls-head-forward-and-down/*noseband type: more horses *Pull-head-forward-and-down* in other noseband than expected (SR = 1.52), and fewer horses *Pull-head-forward-and-down* in a Cavesson than expected (SR = -1.13).

Figure 2. Significant associations ($P < 0.05$) between identified behavior and single items of apparatus. Graphs depict positive and negative standardised residual (SR) values for each significant behavior/apparatus category, indicating the direction of association for each. Behaviors are: Figure2a) *Loves-to-jump-obstacles*, Figure2b) *Needs-strong-cues-to-canter*, Figure3c) *Pulls-head-forward-and-down* and *Very-sensitive-to-leg-cues*

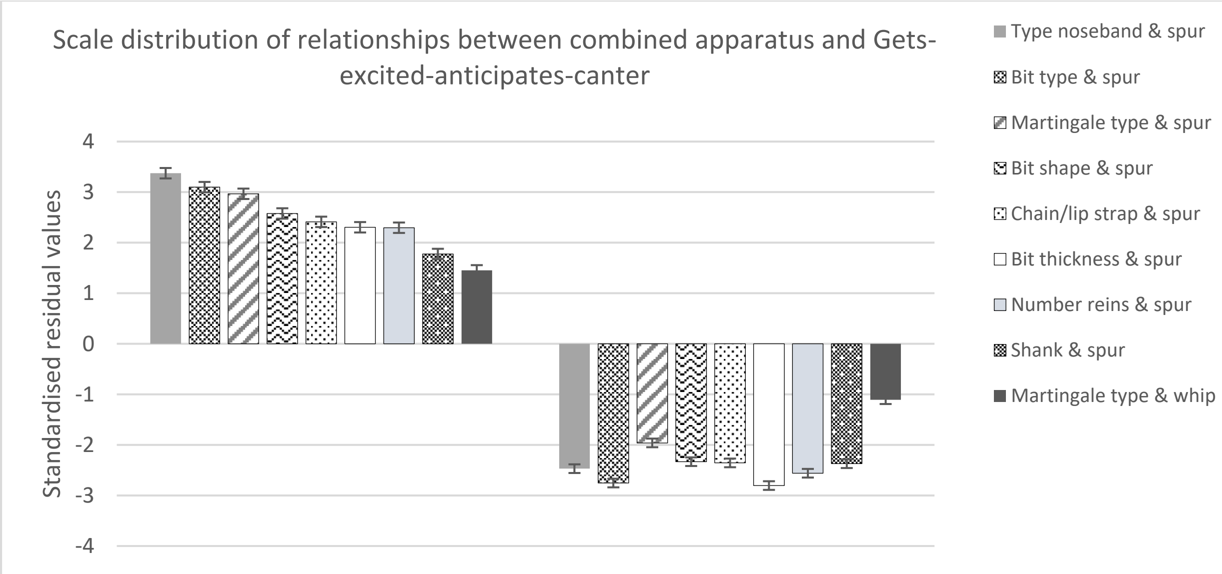
Statistical data analysis - Bitted vs bitless bridles and influence on horse ridden behavior

There were no significant associations between bridle type (i.e. bitted or bitless) and any single unwanted ridden behavior.

Statistical data analysis – associations among behavior and combinations of apparatus responses

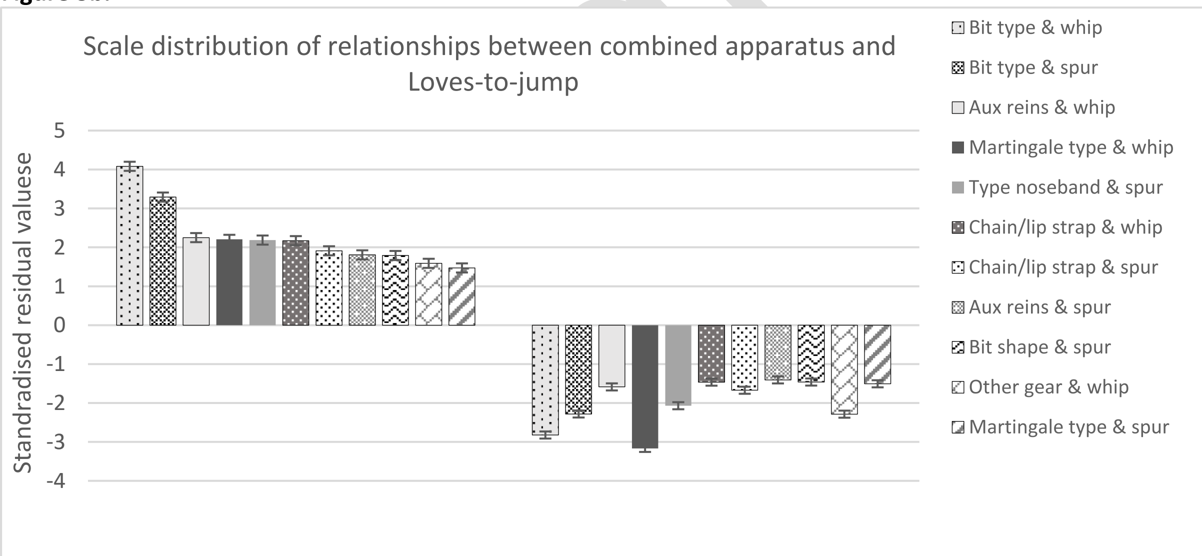
The 121 tables of behavior and concatenated apparatus returned 45 significant results showing that 37.19% of chi-square tests of combinations of apparatus that apply opposing stimuli were significantly associated with certain behavioral reports. Significant results are shown in Figure 3. A copy of statistical data analysis in tabulated form is available as supplementary material.

Figure 3a.



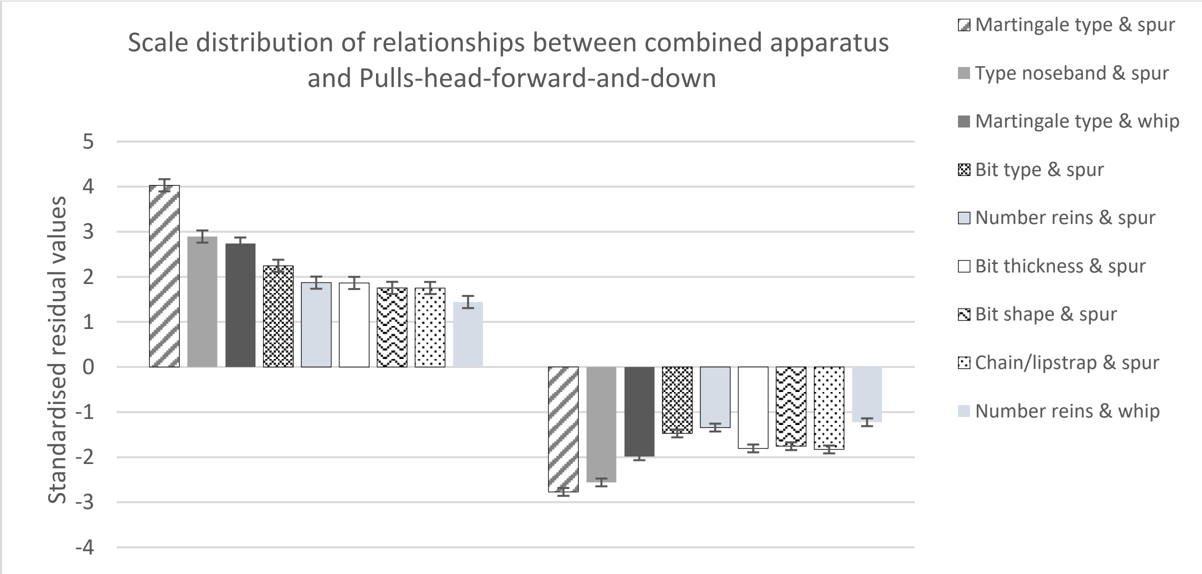
Example, *Gets-excited-anticipates-canter*/bit type & spur: more horses *Get-excited-anticipates-canter* with snaffle only (no spurs) than expected (SR = 3.10) and fewer horses *Get-excited-anticipates-canter* with snaffle & spurs than expected (SR = -2.75)

Figure 3b.



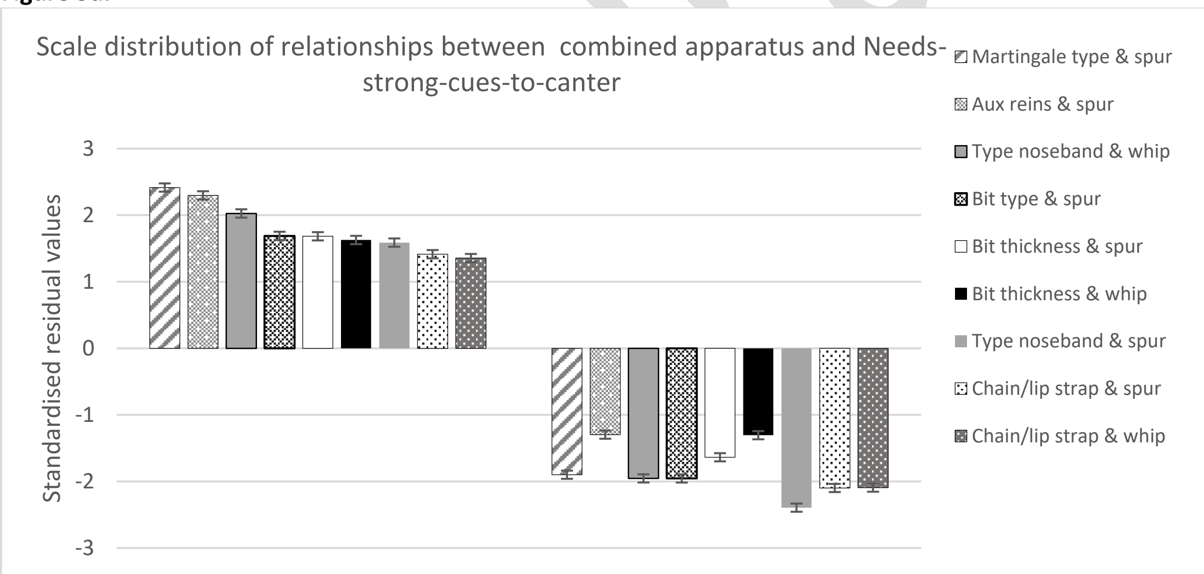
Example, *Loves-to-jump*/martingale type & spur: more horses do not *Love-to-jump* with standing martingale & spurs than expected (SR = 1.47) and fewer horses do not *Love-to-jump* with other martingale & spurs than expected (SR = -1.51)

Figure 3c.



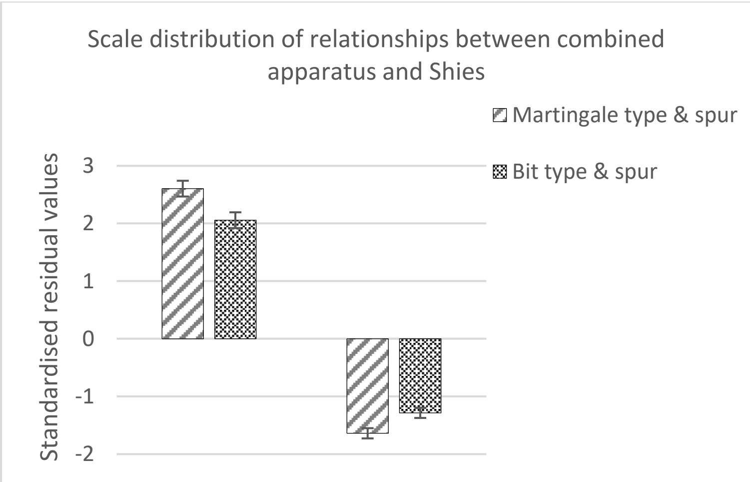
Example, *Pulls-head-forward-and-down/*bit thickness & spur*:* more horses *Pull-head-forward-and-down* with medium thickness bit only (no spurs) than expected (SR = 1.86) and fewer horses *Pull-head-forward-and-down* with a medium thickness bit & spurs than expected (SR = -1.81)

Figure 3d.



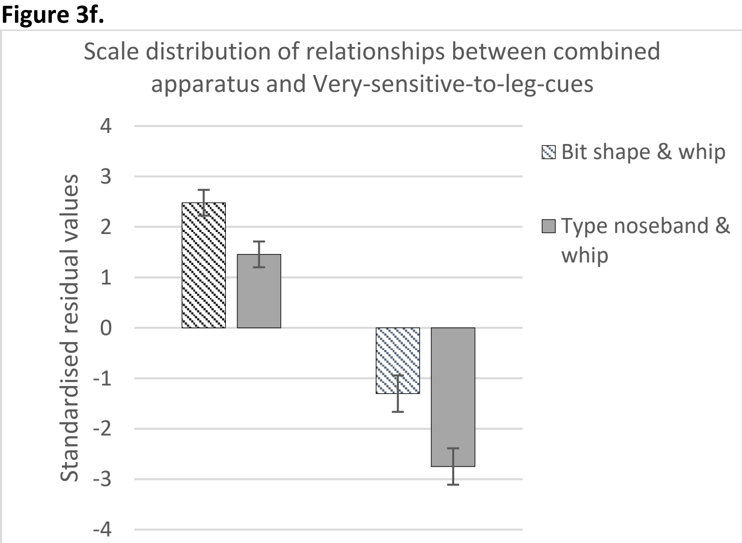
Example, *Needs-strong-cues-to-canter*/bit thickness & whip: more horses *Need-strong- cues-to canter* with thin bit & whip than expected (SR = 1.63) and fewer horses *Need-strong-cues-to-canter* with only a bit (no whip) than expected (SR = -1.31)

Figure 3e.



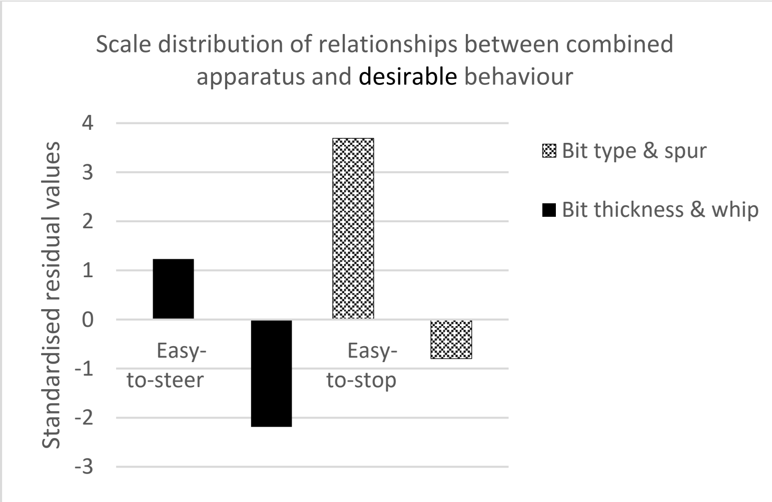
Example, *Shies*/martingale type & spur: more horses *Shy* with martingale only (no spurs) than expected (SR = 2.60) and fewer horses *Shy* with spurs only (no martingale) than expected (SR = -1.64)

Figure 3f.



Example, *Very-sensitive-to-leg-cues/*bit shape & whip: more horses are not *Very-sensitive-to-leg-cues* with complex (roller, key etc) bit & whip than expected (SR = 2.48) and fewer horses are Very-sensitive-to-leg-cues with complex (roller, key etc) bit & whip than expected (SR = -1.30)

Figure 3g.



Example, *Easy-to-steer*/bit thickness & whip: more horses are not *Easy-to-steer* with only one aversive than expected (SR = 1.23) and fewer horses are not *Easy-to-steer* with medium thickness bit & whip than expected (SR = -2.19)

Figure 3. Significant associations ($P < 0.05$) between ridden horse behavior and combinations of apparatus that apply opposing stimuli. Graphs depict positive and negative standardised residual (SR) values for each significant behavior/apparatus combination, indicating the direction of association for each.
Behaviors are: Figure 3a) *Gets-excited-anticipates-canter*, Figure 3b) *Loves-to-jump*, Figure 3c) *Pulls-head-forward-and-down,* Figure 3d) *Needs-strong-cues-to-canter,* Figure 3e) *Shies,* Figure3f) *Very-sensitive-to-leg-cues,* Figure 3g) *Easy-to-steer; Easy-to-stop*

The behavior *pulls-head-forward-and-down* in association with combinations of apparatus generated 9 significant associations with items of gear in combination: 7 when decelerating apparatus (bit type, bit thickness, bit shape, number of reins, lip strap/chain, type noseband and martingale type) were combined with spurs and 2 when decelerating apparatus (martingale type and number of reins) were combined with whips. Equine responses to riders' leg cues generated 12 significant results. *Very-sensitive-to-leg-cues* showed significant associations with both noseband type and bit mouthpiece type (decelerating devices) when these were combined with whips (accelerating devices). In contrast, the behavior *needs-strong-canter-cues* generated 10 significant associations: 6 when items of decelerating apparatus were combined with spurs and 4 when items of decelerating apparatus were combined with whips.

The behavior *gets-excited-anticipates-canter* generated 9 significant associations with items of gear in combination: 8 when items of decelerating apparatus were combined with spurs and 1 when items of decelerating apparatus were combined with whips. *Shying* showed two significant associations with spur use: one relating to spur and martingale use and one relating to spur and bit type. *Ease-of-steering* was associated with combined bit thickness and whip use and *ease-of-stopping* was associated with bit style and spur combinations. *Loves-to-jump-obstacles* showed 11 significant associations: 6 when items of decelerating apparatus were combined with spurs and 5 when decelerating apparatus were combined with whips.

Multivariate analysis

Model A: Very-sensitive-to-leg-cues

Type of head gear ($P = 0.01$) and noseband type and whip ($P = 0.05$) were significantly associated with increased sensitivity to leg cues. Horses wearing a plain cavesson or other forms of head gear were 3.8 times ($P = 0.009$; CI: 1.39 – 10.40) and 4.5 times ($P = 0.002$; CI: 1.69-11.76) more likely to be sensitive to the leg cues than those wearing headcollars or a halter. Horses wearing another head gear type and whip were 4.4 times more likely ($P = 0.02$; CI: 1.30-14.56) to be extra sensitive to leg cues than horses in headcollar / halters. Hosmer Lemeshow goodness-of-fit statistics confirmed that the model showed a good fit ($P > 0.05$). The likelihood ratio chi-square test of model coefficients revealed a significance level of $P \leq 0.05$ at each step. ROC curve analysis indicated that the predictability of the final model to determine risk factors associated with decreased sensitivity to the leg cues was average (ROC: 0.52).

Model B: Needs-strong-canter-cues

Horses ridden with smooth spurs were 3.07 times more likely ($P = 0.005$; CI: 1.41-6.68) to require stronger canter cues than those not ridden with spurs. The presence of other auxiliary equipment was significantly associated with horses requiring stronger canter cues ($P = 0.002$); horses ridden in a martingale were 0.17 times less likely ($P = 0.0004$; CI: 0.06-0.45) to require stronger canter cues than horses ridden with spurs. Hosmer Lemeshow goodness-of-fit statistics ($P > 0.05$) and likelihood ratio chi-square test of model coefficients ($P < 0.05$) confirmed that the model showed a good fit. ROC curve analysis indicated that the predictability of the final model to determine risk factors associated with reduced canter cues was moderate (ROC: 0.62).

Model C: Gets-excited-anticipates-canter

Horses ridden with spurs and a bit with a shank were 1.6 times more likely to anticipate canter than horses ridden in a bit with no shank and spurs (P = 0.04; CI: 1.01-1.71). Horses wearing a martingale only or wearing a martingale and ridden with spurs were 8.5 times (P = 0.03; CI: 1.20-60.44) and 8.6 (P = 0.013; CI: 1.57-47.42)] more likely, respectively, to anticipate canter than horses ridden in spurs alone. The use of other auxiliary equipment was significantly associated with increased anticipation of canter (P = 0.002), with horses ridden in a martingale or other equipment (draw reins, side reins, Whitaker reins, Pessoa® Training System, Market Harborough, tie-down, Chambon) being less likely to anticipate canter than horses ridden with spurs, only (0.06 times, P = 0.003, CI: 0.10-0.38; 0.21 times, P = 0.0004; CIs: 0.09-0.50, respectively). Hosmer Lemeshow goodness-of-fit statistics ($P > 0.05$) and likelihood ratio chi-square test of model coefficients ($P < 0.05$) confirmed that the model showed a good fit. ROC curve analysis indicated that the predictability of the final model to determine risk factors associated with reduced anticipation to canter was moderate (ROC: 0.66).

Model D: Easy-to-stop

Horses ridden in a bit and with spurs were 1.9 times more likely to stop easily than horses ridden in a bit alone (P = 0.05; CI: 1.00-3.62). Meanwhile horses ridden in double reins were 0.28 times less likely to stop easily than horses ridden in single reins (P = 0.03; CI: 0.08-0.92). Hosmer Lemeshow goodness-of-fit statistics ($P > 0.05$) and likelihood ratio chi-square test of model coefficients ($P < 0.05$) confirmed that the model showed a good fit. ROC curve analysis indicated that the predictability of the final model to determine risk factors associated with reduced anticipation to canter was moderate (ROC: 0.68).

Model E: Loves-to-jump-obstacles

Using a martingale type and a whip were significantly associated with horses who were reported as loving to jump obstacles (P = 0.002); whereas horses ridden with just a whip were 0.2 times less likely to be reported as loving to jump than those ridden in a standing martingale (P = 0.004; CI: 0.064 to 0.602). Although associations between the reported love of jumping and noseband-and-spur use in combination were non-significant ($P = 0.09$), the combination was retained as it improved model fit. Hosmer Lemeshow goodness-of-fit statistics ($P > 0.05$) and likelihood ratio chi-square test of model coefficients ($P < 0.05$) confirmed that the model showed a good fit. ROC curve analysis indicated that the predictability of the final model to determine risk factors associated with reduced anticipation to canter was moderate (ROC: 0.65).

Discussion

Due to the omnipresence of NR via physical apparatus in horse riding and handling, it is critical that equestrians recognise the potential aversive nature of pressure-applying devices such as bits, reins, whips and spurs and learn to use them optimally. This study of Australian horse-owners' use of gear and their reports of their horses' behavior permitted an exploration of associations among apparatus commonly used in equitation and the occurrence of apparent conflict behavior whilst being ridden. Eleven horse behaviors exhibited when ridden were considered. Of these, two behaviors, *easy-to-steer* and *easy-to-stop* are typically desirable in the ridden horse. The remaining nine are behaviors commonly considered undesirable in equitation, including three, *rearing, bolting* and *bucking,* considered extreme displays of conflict (McGreevy et al., 2018b). While not inferring causation, the current results reveal correlational associations among commonly used apparatus and some ridden behavior. The study has shown that the use of numerous combinations of both accelerator (whips and spurs) and decelerator devices (bits, reins, chain/lip strap, nosebands, martingales and auxiliary equipment), both in combination and separately, are associated with various unwelcome behaviors in horses. Multivariate modeling also suggests that choice of apparatus may indeed reflect rider perception of the horses' ridden responses. For example, a less reactive horse may prompt the rider to select a whip or spurs to motivate locomotory response.

Whilst the expression of ridden behavior cannot be specifically linked to a single item of apparatus, multiple items of apparatus applying opposing stimuli used in conjunction can influence behavior in the ridden horse. The significant relationships among them will be discussed separately.

Ridden horse behavior

The current study revealed behaviors (n=9) that may be indicative of discomfort, conflict or habituation (McGreevy et al., 2018b) and potentially reflect welfare issues. For example, the behaviors *very-sensitive-to-leg-cues* and *gets-excited-and-anticipates-canter* may be manifestations of hyper-reactivity and may indicate that the horses affected are hyper-vigilant to pressure cues on the thorax or abdomen. Horses that show anticipation of accelerator cues may motivate their riders to select strong decelerator cues to match. The combination of strong accelerator and strong decelerator devices may detract from the horse's *telos* and may contribute to learned helplessness (Hall et al., 2008). *Loves-to-jump-obstacles* may be used to euphemistically describe horses that rush fences, and so horses described this way may be showing a hyper-reactive response (McGreevy et al., 2018b). *Shying* is also a hyper-reactive response reflecting tendencies to favour turning to one direction (McGreevy, 2012; McLean and McLean, 2008), with *ease-of-steering* being a manifestation of the converse (McLean and McLean, 2008). The behavior *needs-strong-canter-cues* may be an indication of habituation to accelerating stimuli (McGreevy, 2012). Interestingly 78.45% of respondents described their horses as *very-sensitive-to-leg-cues*. That said, riders' perception of horse sensitivity may be subjective. Some riders consider this a desirable trait and so observations of this trait may be affected by reporting bias.

A less subjective behavior, *pull-head-forward-and-down*, was reported by 30.74% of respondents. This may be of concern as this behavior could be a flawed response to rider deceleration cues and/or restriction of the mouth/head/neck and is certainly classed as an unwelcome response to stimuli (McGreevy et al., 2018b).

Interestingly, behaviors considered to be extreme displays of conflict, *rearing, bolting* and *bucking* were reported to occur at least sometimes by 3.31%, 3.41% and 7.34%, respectively. As these three conflict behaviors are potentially dangerous, animals displaying them tend to not make desired riding horses (Hawson et al., 2011; Oddie et al., 2014; McGreevy et al., 2015) and may be culled from the more general riding horse population (Thomson et al., 2014). Culling of horses with dangerous traits may account for the relatively low numbers in this survey and should be considered when interpreting results.

Apparatus used

The number of apparatus options offered in the original survey was extensive. This is an authentic reflection of the diverse choices possible in the saddlery market (Doherty et al., 2017b). For example, the categories for bit composition (n=57) and noseband type (n=35) indicate the vast array of apparatus available to the modern equestrian. Choice of apparatus is influenced by discipline (Hill et al., 2015) and by other multiple factors, such as fashion and peer pressure, relating to the rider (Williams and Tabor, 2017). However, there remains only limited scientifically validated information available to help riders make informed choices when selecting bits for their horses. Subjective opinions among riders have been proposed as the main determinant in bit selection, but equine oral anatomy, individual horse sensitivity and past experiences should be also be considered (Manfredi et al., 2007). The naïve horse seems to find oral pressures from bits aversive but may habituate quickly (Christensen et al., 2011). Over the course of a working life, the cumulative effect on horses of strong rein tension and a restrictive noseband may magnify bit pressure, cause discomfort and/or pain consequently affecting behavior and compromising welfare (McGreevy et al., 2012; Weller et al., 2020).

Manfredi et al. (2009) used lateral fluoroscopic imaging to investigate intra-oral behaviors in six horses wearing three different snaffle bits, with and without rein tension of 25N. Behaviors indicative of discomfort, such as resistance to stopping and turning and pulling-head-forward-and-down, are readily apparent to most informed observers (Dyson and Pollard, 2020) while other subtle intra-oral behaviors that are not observable externally, may also indicate discomfort and/or stress. Manfredi et al. (2009) detected significant increases in tongue movements in response to static bilateral rein tension of 25 ± 5N but behavioral responses were specific to individual horses, rather than to bit type. A large variety of bits are available to riders, with many marketed towards resolution of oral, behavioral and training problems (McGreevy et al., 2014). These anecdotal beliefs may account for the high degree of variation found in the current respondents' answers to bit-specific questions.

Snaffle bits were the predominant reported bit type used. This was expected as snaffles are a basic bit design but the reported use of these bits incorporated many variations in mouthpiece style, composition and thickness, and therefore severity.

Ridden horse behavior in relation to single items of apparatus

In the current study preliminary analysis to identify significant relationships between behavior and apparatus using broad category parameters exposed no associations between any behavior/single apparatus combinations. Secondary analysis revealed 13 significant associations between *ridden horse behavior and single items of apparatus*. Of these, 4 (30.77%) were between apparatus and the behavior *needs-strong-canter-cues*. Two of these relationships involved bit type and chain/lip strap and the other 2 involved the accelerating devices, whips and spurs. This suggests that riders used these devices to address issues of habituation and to augment signals from their legs. There were 7 significant associations (53.85%) between apparatus and the behavior *loves-to-jump-obstacles.* This suggests that the behavior labelled *loves-to-jump-obstacles* has more than one cause and its origins may be obscured by subjective interpretations. It also indicates that the relationships among gear and appearing to enjoy jumping are especially complex. It will be important that further studies of this perceived attribute trait are designed to unpick its different manifestations.

Bitted versus bitless bridles and horse ridden behavior

Whilst it may appear convenient that the equine mouth can accommodate a bit, allowing for directional and speed control via pressure, it did not evolve to accommodate a bit nor to facilitate NR (McGreevy et al., 2014). Bitless bridles appear to be gaining popularity, particularly among leisure and endurance riders, whose choice of apparatus is not restricted by current competition rules (Hill et al., 2015). Riders may choose this equipment as they perceive the bitless bridle to be less aversive to the horse than bitted bridles or they may not feel the need to ride with the constraints currently required in many competition arenas. There is also the more recent view that bits may interfere with the natural seal of the horse’s mouth during respiration, affecting athletic performance (Mellor and Beausoleil, 2017). The design of bitless bridles varies, as does their ability to apply pressure to sensitive structures of the head/nose and their degree of aversiveness. For example, even though it is bitless, the hackamore bridle incorporates lever action to apply amplified, potentially severe, rein pressure to the nasal bones and mandible. In contrast, the bitted bridle applies pressure to sensitive oral structures of the mouth; the hard palate, tongue, bars and lips. The potential deleterious impact of any bridle on the horse is therefore not only influenced by its design but also its operation. Any bit style improperly fitted or used in ridden horses can potentially cause oral trauma such as bone spurs, ulceration, and tongue lacerations (Mata et al., 2015).

The current indirect investigation into the anecdotal assumption that bitless bridles are less aversive than bitted bridles revealed interesting results. Analysis found the use of bitted or bitless bridles had no significant association with the prevalence of conflict behaviors. Scofield and Randle (2013) compared the effects of bitted and bitless bridles on behavior in a pilot study. Results suggested the occurrence of conflict behavior was similar with both types of equipment and the current study appears to support these findings. However, in the absence of specific details of bitless bridle design and their corresponding mechanisms, further investigations are warranted.

Behavior versus combinations of equipment

The current study revealed a series of significant relationships between behavior and the use of items of apparatus. Items that deliver opposing stimuli revealed differing results to that of single apparatus use. Applying the same behavior and apparatus categories used for single apparatus, apparatus designed to amplify pressures for acceleration or deceleration were then combined and 37.19% of these relationships were significant. The use of items of apparatus that apply opposing stimuli in combination may be both counterproductive and unethical (McLean and McGreevy, 2010a).

Habituation is not the only potential consequence of the use of simultaneous, opposing cues. Horses may become either difficult to stop or sensitive to leg signals, therefore, cues applied as light versions of the original stimulus, and thus horsemanship in general, are compromised. All other conflict or discomfort behaviors in response to opposing stimuli that the horse cannot resolve reflect serious insults to welfare.

It should be recognised that nosebands and bits are not used in isolation in the ridden horse. Previous studies have suggested wearing a noseband or bit in the absence of rein tension, regardless of bit type, is sufficient to produce a stress response in naïve horses (Manfredi et al., 2009; McGreevy et al., 2012). By design, restrictive nosebands inhibit oral movements (Fenner et al., 2016; Doherty et al., 2017a). In a bid to find comfort, horses fight against this constraint; a response that may cause the structural changes in nasal bones and mandibles that have recently been reported at the site of restrictive nosebands (Perez-Manrique et al., 2020). There is evidence that the additional impact of rein tension and other auxiliary equipment, such as martingales and draw reins, may further amplify pressures applied by the bridle, noseband and bit (McGreevy et al., 2014). The additional complications of the physical influence of the rider on the horse, rider skill level and experience and decisions affecting apparatus choice, further confound the effect of apparatus upon horse behavior (Williams and Tabor, 2017).

Restrictive apparatus applied to the mouth and other parts of the head is primarily used to increase control and to steer forward movement using pressure; and therefore to influence deceleration through negative reinforcement. In contrast, whips and spurs are generally used to augment the cues of the riders' legs to promote and achieve acceleration. The separate use of accelerating and decelerating apparatus has been the focus of many previous studies (Heleski et al., 2009; McLean and McGreevy, 2010a; Clayton et al., 2011; Doherty et al., 2017a; McGreevy et al., 2017) but studying the concurrent use of two such opposing types of stimuli and the resultant effect upon behavior is novel. Overuse of any apparatus that directs pressure onto animals can lead to habituation (McGreevy and Boakes, 2011) and may easily become a welfare concern if, instead of improving their technique and most notably their timing, riders escalate pressure (e.g., by using more severe apparatus) to elicit desired behavior. Combined with the potential for intentional or accidental misuse by the rider, the use of apparatus to apply opposing tactile stimuli to the ridden horse may result in significant welfare compromises. The study indicates that no single item of apparatus is exclusively associated with the expression of behavior issues in horses. However, it goes on to reveal associations among behaviors and the apparatus issuing opposing aversive stimuli. This reminds us that one must consider numerous factors when evaluating the aetiology of equine responses including rider application of stimuli, horse sensitivity and general context.

Limitations

Cross-sectional surveys, such as the current one, cannot reliably identify casual relationships. Additionally, differences in horse age, sex, learning history, and coping style may influence the responses that emerge under saddle, and data on these were not collected by the current survey instrument. The anonymity of internet surveys promotes self-disclosure and may consequently negate the tendency for respondents to provide socially acceptable responses (Fricker and Schonlau, 2002). Regardless, despite appropriate aspects of survey design and administration to minimise sanitised reporting, the prospect that some respondents offered responses perceived as desirable must still be noted as a potential limitation.

In the absence of a comprehensive ridden behavior ethogram to facilitate accurate behavioral definition, the current responses may have been skewed as a function of respondents' interpretation of their horses' behavior (Pierard et al., 2019). Respondents' interpretation of some behaviors may reflect subjectivity and anthropomorphism, and this should be considered when interpreting the current results.

Riding horses displaying extreme displays of conflict, including *rearing, bolting* and *bucking* may be misrepresented in this study. Further investigation into longitudinal associations between hazardous behaviors and apparatus is warranted. Additionally, for the purpose of interpretation, the availability of an unambiguous, detailed ethogram to enable the recording of behavioral observations in relation to behavior indicative of a state of conflict would assist the evaluation of apparatus, training and rider influence and allow comprehensive inferences to be drawn about the welfare of the horse (Hall and Heleski, 2017).

Raw data analyzed in this study were collected by Hill et al. (2015). It should be considered that responses may differ over time, and responses to a contemporary survey may yield different results.

Conclusions

The current study indicates that the use of a single specific item of apparatus cannot be fully implicated in the expression or intensity of undesirable ridden horse behavior. Similarly, the use of bitted or bitless bridle was shown to have no significant effect on the reports of ridden horse behavior. However, significant non-causal associations were demonstrated among behaviors and apparatus that imposes opposing aversive stimuli. This suggests multiple items of apparatus can influence behavior when used simultaneously, particularly if these devices, when used in combination, amplify contradictory stimuli simultaneously. Further investigation is required to identify all factors involved in the ontogeny and expression of conflict behavior in ridden horses.

Acknowledgements

The authors thank the respondents to this survey and Hill et al. (2015) for access to the raw data.

Conflicts of Interest

The authors declare no conflicts of interest.

References

Anthony, D.W., Brown, D.R., 2015. The origins of horseback riding. Antiquity 65**,** 22-38.

Bailey, C., Rose, R., Reid, S., Hodgson, D., 1997. Wastage in the Australian thoroughbred racing industry: a survey of Sydney trainers. Aust. Vet. J. 75(1), 64-66.

Christensen, J.W., Zharkikh, T.L., Antoine, A., Malmkvist, J., 2011. Rein tension acceptance in young horses in a voluntary test situation. Equine Vet. J. 43**,** 223-228.

Clayton, H.M., Larson, B., Kaiser, L.J., Lavagnino, M., 2011. Length and elasticity of side reins affect rein tension at trot. Vet. J. 188**,** 291-294.

Cook, W., Mills, D., 2009. Preliminary study of jointed snaffle vs. crossunder bitless bridles: Quantified comparison of behaviour in four horses. Equine Vet. J. 41**,** 827-830.

Crago, F., Shea, G., James, O., Schemann, K., McGreevy, P.D., 2019. An opportunistic pilot study of radiographs of equine nasal bones at the usual site of nosebands. J. Vet. Behav.: Clin. Appl. Res. 29**,** 70-76.

Cross, G.H., Cheung, M.K.P., Honey, T.J., Pau, M.K., Senior, K.J., 2017. Application of a dual force sensor system to characterize the intrinsic operation of horse bridles and bits. J. Equine Vet. Sci. 48**,** 129-135.e3.

Doherty, O., Casey, V., McGreevy, P., Arkins, S., 2017a. Noseband use in equestrian sports – an international study. PLoS One 12**,** e0169060.

Doherty, O., Casey, V., McGreevy, P., Mclean, A., Parker, P., Arkins, S., 2017b. An analysis of visible patterns of horse bit wear. J. Vet. Behav.: Clin. Appl. Res. 18**,** 84-91.

Doherty, O., McGreevy, P.D., Pearson, G., 2017c. The importance of learning theory and equitation science to the veterinarian. Appl. Anim. Behav. Sci 190**,** 111-122.

Dyson, S., Pollard, D., 2020. Application of a ridden horse pain ethogram and its relationship with gait in a convenience sample of 60 riding horses. Animals 10**,** 1044.

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.

Fricker, R.D., Schonlau, M., 2002. Advantages and disadvantages of internet research surveys: evidence from the literature. Field Methods 14**,** 347-367.

Hall, C., Goodwin, D., Heleski, C., Randle, H., Waran, N., 2008. Is there evidence of learned helplessness in horses? J. Appl. Anim. Welf. Sci. 11**,** 249-266.

Hall, C., Heleski, C., 2017. The role of the ethogram in equitation science. Appl. Anim. Behav. Sci. 190**,** 102-110.

Hall, C., Huws, N., White, C., Taylor, E., Owen, H., McGreevy, P., 2013. Assessment of ridden horse behavior. J. Vet. Behav.: Clin. Appl. Res. 8**,** 62-73.

Hausberger, M., Bruderer, C., Scolan, N.L., Pierre, J.-S., 2004. Interplay between environmental and genetic factors in temperament/personality traits in horses (*Equus caballus*). J. Comp. Psych. 118**,** 434-446.

Hausberger, M., Muller, C., 2002. A brief note on some possible factors involved in the reactions of horses to humans. Appl. Anim. Behav. Sci. 76**,** 339-344.

Hawson, L.A., Oddie, C., McLean, A.N., McGreevy, P.D., 2011. Is safety valued in the Australian pony market? J. Vet. Behav.: Clin. Appl. Res. 6**,** 254-260.

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.

Hill, E., McGreevy, P.D., Caspar, G., White, P., McLean, A.N., 2015. Apparatus use in popular equestrian disciplines in Australia. J. Vet. Behav.: Clin. Appl. Res. 10**,** 147-152.

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(1), 15-19.

Jones, B., McGreevy, P.D., 2010. Ethical equitation: Applying a cost-benefit approach. J. Vet. Behav.: Clin. Appl. Res.5**,** 196-202.

Lescolan, N., Hausberger, M., Wolff, A., 1997. Stability over situations in temperamental traits of horses as revealed by experimental and scoring approaches. Behav. Process 41, 257-266.

Levine, M.A., Bailey, G., Whitewell, K.E., Jeffcott, L.B., 2000. Palaeopathology and horse domestication: the case of some Iron Age horses horn the Altai Mountains, Siberia. In: Bailey, G., Charles, R., Winder, N., (Eds.), Symposia of the Association for Environmental Archaeology. Oxford: Oxbow Books, pp. 123–133.

Manfredi, J., Clayton, H.M., Rosenstein, D., 2007. Radiographic study of bit position within the horse's oral cavity. Equine. Comp. Ex. Physiol. 2**,** 195-201.

Manfredi, J.M., Rosenstein, D., Lanovaz, J.L., Nauwelaerts, S., Clayton, H.M. 2009. Fluoroscopic study of oral behaviours in response to the presence of a bit and the effects of rein tension. Comp. Ex. Physiol. 6**,** 143-148.

Mata, F., Johnson, C., Bishop, C., 2015. A cross-sectional epidemiological study of prevalence and severity of bit-induced oral trauma in polo ponies and race horses. J. Appl. Anim. Welf. Sci. 18**,** 259-268.

McGreevy, P., 2012. Equine behavior : a guide for veterinarians and equine scientists, Edinburgh, Elsevier, pp. 280-340.

McGreevy, P., Berger, J., De Brauwere, N., Doherty, O., Harrison, A., Fiedler, J., Jones, C., McDonnell, S., McLean, A., Nakonechny, L., 2018a. Using the five domains model to assess the adverse impacts of husbandry, veterinary, and equitation interventions on horse welfare. Animals 8**,** 41.

McGreevy, P., Boakes, R., 2011. Carrots and sticks: principles of animal training, Darlington Press, pp. 63-87.

McGreevy, P., Christensen, J.W., Von Borstel, U.,K., McLean, A., 2018b. Equitation science, John Wiley & Sons, pp. 213-312.

McGreevy, P., McManus, P., 2017. Why horse-racing in Australia needs a social licence to operate. The Conversation, 3. Available at: https://theconversation.com/why-horse-racing-in-australia-needs-a-social-licence-to-operate-79492. Accessed 2nd February, 2020.

McGreevy, P., Oddie, C., Burton, F., McLean, A., 2009. The horse–human dyad: can we align horse training and handling activities with the equid social ethogram? Vet. J. 181**,** 12-18.

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.

McGreevy, P.D., 2007. The advent of equitation science. Vet. J. 174**,** 492-500.

McGreevy, P.D., Doherty O., Channon, W., Kyrklund, K., Webster, J., 2017. The use of nosebands in equitation and the merits of an international equestrian welfare and safety committee: A commentary. Vet. J. 222**,** 36-40.

McGreevy, P.D., Oddie, C.F., Hawson, L.A., McLean, A.N., Evans, D.L., 2015. Do vendors value safety in Thoroughbred horses in the Australian recreational riding horse market? J. Vet. Behav.: Clin. Appl. Res. 10**,** 153-157.

McGreevy, P.D., Sundin, M., Karlsteen, M., Berglin, L., Ternstrom, J., Hawson, L., Richardsson, H., McLean, A.N., 2014. Problems at the human–horse interface and prospects for smart textile solutions. J. Vet. Behav.: Clin. Appl. Res. 9**,** 34-42.

McLean, A.N., Mclean, M.M., 2008. Academic horse training: equitation science in practice. Australian Equine Behaviour Centre, Victoria, Australia, pp. 254-284.

McLean, A.N. Christensen, J.W., 2017. The application of learning theory in horse training. Appl. Anim. Behav. Sci. 190**,** 18-27.

McLean, A.N., McGreevy, P.D., 2010a. Ethical equitation: Capping the price horses pay for human glory. J. Vet. Behav.: Clin. Appl. Res. 5**,** 203-209.

McLean, A.N., McGreevy, P.D., 2010b. Horse-training techniques that may defy the principles of learning theory and compromise welfare. J. Vet. Behav.: Clin. Appl. Res. 5**,** 187-195.

Mellor, D.J., Beausoleil, N.J., 2017. Equine welfare during exercise: An evaluation of breathing, breathlessness and bridles. Animals 7**,** 41.

Mills, D.S., Marchant-Forde, J.N., 2010. The encyclopedia of applied animal behaviour and welfare, CABI. p. 127.

Moberg, G., M, J.A., 2000. Biology of animal stress : basic principles and implications for animal welfare, Wallingford, UNITED KINGDOM, CABI, pp. 77-78.

Oddie, C.F., Hawson, L.A., McLean, A.N., McGreevy, P.D., 2014. Do vendors value safety in the Australian recreational (non-thoroughbred) riding horse market? J. Vet. Behav.: Clin. Appl. Res. 9**,** 375-381.

Parkin, T., Clegg, P., French, N., Proudman, C., Riggs, C., Singer, E., Webbon, P., Morgan, K., 2006. Analysis of horse race videos to identify intra-race risk factors for fatal distal limb fracture. Prev. Vet. Med. 74, 44-55.
Pierard, M., McGreevy, P., Geers, R., 2019. Developing a descriptive reference ethogram for Equitation Science. J. Vet. Behav: Clin. Appl. Res. 29**,** 148.
Randle, H., Steenbergen, M., Roberts, K., Hemmings, A., 2017. The use of the technology in equitation science: A panacea or abductive science? Appl. Anim.Behav. Sci. 190**,** 57-73.
Randle, H.,Waran, N., 2017. Breaking down barriers and dispelling myths: The need for a scientific approach to Equitation. Appl. Anim. Behav. Sci. 190**,** 1-4.
Randle, H., Waran, N., 2019. Equitation Science in Practice: how collaboration, communication and change can improve equine welfare. J. Vet. Behav.: Clin. Appl. Res. 29**,** viii-x.
Scofiel, R., Randle, H., 2013. Preliminary comparison of behaviors exhibited by horses ridden in bitted and bitless bridles. J. Vet. Behav.: Clin. Appl. Res. 8**,** e20-e21.
Smith, L.J., Tabor, G., Williams, J., 2018. A retrospective case control study to investigate race level risk factors associated with horse falls in Irish point-to-point races. Comp. Ex. Phys. 14(2), 27-134.
Søndergaard, E., Ladewig, J., 2004. Group housing exerts a positive effect on the behaviour of young horses during training. Appl. Anim. Behav. Sci. 87**,** 105-118.
Thomson, P., Hayek, A., Jones, B., Evans, D., McGreevy, P., 2014. Number, causes and destinations of horses leaving the Australian thoroughbred and standardbred racing industries. Aust. Vet. J. 92**,** 303-311.
Uldahl, M., Clayton, H., 2019. Lesions associated with the use of bits, nosebands, spurs and whips in Danish competition horses. Eq. Vet. J. 51**,** 154-162.
Weller, D., Franklin, S., Shea, G., White, P., Fenner, K., Wilson, B., Wilkins, C., McGreevy, P., 2020. The reported use of nosebands in racing and equestrian pursuits. Animals 10**,** 776.
Williams, J.M., Marlin, D.M., Langley, N., Parkin, T.D., Randle, H., 2013. The Grand National: a review of factors associated with non-completion and horse-falls, 1990 to 2012. Comp. Ex. Phys. 9(3-4), 131-146.
Williams, J., Tabor, G., 2017. Rider impacts on equitation. Appl. Anim. Behav. Sci. 190**,** 28-42.

Appendices