

Evolving Protocols for Research in Equitation Science

Marc Pierard^{a*}, Carol Hall^{b*}, Uta König von Borstel^{c*}, Alison Averis^{d*}, Lesley Hawson^e, Andrew McLean^f, Charlotte Nevison^g, Kathalijne Visser^h, Paul McGreevy^e

^aDepartment of Biosystems, KU Leuven, Bijzondere Weg 12, 3360 Lovenjoel, Belgium

^bSchool of Animal, Rural & Environmental Sciences, Nottingham Trent University, Brackenhurst Campus, Southwell, Nottinghamshire, NG25 0QF, UK

^cDepartment of Animal Breeding and Genetics, University of Goettingen, Albrecht-Thaer-Weg 3, 37075, Goettingen, Germany

^d6 Stonelaws Cottages, East Linton, East Lothian, EH40 3DX, UK

^eFaculty of Veterinary Science, University of Sydney, NSW 2006, Australia

^fAustralian Equine Behaviour Centre, Broadford, VIC 3658, Australia

^gDepartment of Life Sciences, Anglia Ruskin University, Cambridge, UK

^hHorsonality, Skipper 3, 8456 JB De Knipe, The Netherlands

*equal first authors

Abstract

Within the emerging discipline of Equitation Science the application of consistent methodology, including robust objective measures, is required for sound scientific evaluation. This report aims to provide an evaluation of current methodology and to propose some initial guidelines for future research. The value of research, especially that involving small sample sizes, can be enhanced by the application of consistent methodology and reporting enabling results to be compared across studies. This paper includes guidelines for experimental design in studies involving the ridden horse. Equine ethograms currently used are reviewed and factors to be considered in the development of a ridden-horse ethogram are evaluated. An assessment of methods used to collect behavioral and physiological data is included and the use of equipment for measurements (e.g., rein-tension and pressure-sensing instruments) is discussed. Equitation science is a new discipline, subject to evolving viewpoints on research foci and design. Technological advances may improve the accuracy and detail of measurements, but must be used within appropriate and valid experimental designs.

Keywords: equitation science; stress; behavior, equine; welfare; ethogram; horse-human interaction

Introduction

Equitation science seeks to evaluate the horse-human partnership and gather scientific evidence that can improve the training, performance and welfare of the ridden horse. Equitation science incorporates multi-disciplinary methodologies, including approaches taken from anatomy and physiology, ethology, veterinary medicine, physics and psychology. To date there have been a number of studies set up to evaluate specific factors that may affect the ridden horse, including tack and equipment (e.g. Manfredi et al., 2005; Heleski et al., 2009; Quick and Warren-Smith, 2009), training (McLean, 2008; Egenvall et al., 2012) and the rider (König von Borstel et al., 2011a). Such studies form the basis for the integrated discipline of equitation science. However methodological problems and small sample sizes continue to adversely affect some aspects of ridden-horse research. Standardization of management procedures is often difficult because horses belong to different owners. Riders, handlers or owners are often part of the research and constitute an additional source of uncontrolled variation. The development and use of robust, valid objective measures will encourage data collection and appropriate methodological reporting that will advance the science of equestrianism for the potential benefit of both horses and humans.

Due to the way horses are owned and/or managed and the nature of most research, some problems commonly occur in experimental design. We outline some basic principles of experimental design to be considered when planning a study involving ridden horses and include recommendations for the collection/recording of behavioral and physiological data. We provide guidelines for studies that assess learning and training. Within the section on experimental design we include recommendations for factors that should be recorded (sample variables, rider characteristics and environmental factors). The development of a ridden-horse ethogram is the first step in quantifying observations of behavioral responses of horses during riding. The basis of equitation science lies in the development of ethograms with reliable and valid parameters enabling a consistent and standardized assessment of the horse's response. Technology exists that allows us to measure some interactions between rider and horse (e.g., rein-tension gauges and pressure sensors), so recommendations for its use, and its limitations, are discussed.

We hope that this report will be the first stage in the standardization of equitation science methods. That will allow more rapid progress within this discipline and facilitate the collection of evidence. Readers are strongly encouraged to visit the research pages of the International Society for Equitation Science's (ISES) website (www.equitationsscience.com) to check for updates.

The design of experiments involving ridden horses and reporting results

Ethical considerations and informed consent

The effect of studies on the horses and the justification for/mitigation of any residual distress is the focus of animal ethics committees. The work must be legal under the legal framework of the country in which the study is being conducted. Even observational studies are subject to institutional animal ethics committee review. In developing their studies, researchers must consider the suitability of the

animals studied – careful habituation of the horses to any scientific equipment to be used should be undertaken prior to starting the study. Wherever feasible, veterinary examination of horses prior to their use in studies is strongly advised. Should anyone be concerned about responses to the equipment, horses can be withdrawn from the study and this option must be made clear. Inclusion and exclusion criteria must be stated and followed.

The involvement of human subjects and stakeholders in scientific research also requires ethical review prior to the study. Use of the Physical Activity Readiness Questionnaire (PAR-Q) is encouraged for human participants to indicate potential health concerns. Any human subject must provide written informed consent. This requirement also pertains to owners of the horse(s), the employers of any riders and venue owners. Human participants must be fully informed of all that is required of them during the study. To achieve the goal of producing relevant results, it is important to make the measurement situation as “ecologically valid” as possible. It is likely that both horse and rider will be required to wear scientific equipment and may be asked to perform particular movements, not simply execute their normal training routines. It is likely that further biological information about riders will be recorded (e.g., limb lengths, bodyweight). The type of data that will be taken, and what will happen with it (including issues such as confidentiality of data), should be outlined before the study. Where measurements taken have the potential to reveal health issues in participants, researchers should inform participants of this possibility but also make it clear that measurements are not made in a clinical capacity and are not diagnostic. Researchers should consider how they will advise the participant if such situations arise. Health and safety concerns must be incorporated into any study, which is required as part of an ethical review process of any academic institution.

Human participants retain the right to withdraw themselves and their data from the study up to a specified time point, typically the point of submission for publication. Unless explicit consent is given, reporting of results must be in an anonymous format.

As equine behavioral medicine matures, we will become better at identifying individuals at high risk for certain disorders. Evidence of resistance to extinction as an attribute of stereotypic horses (e.g., Parker et al., 2008) suggests that we should also become more skilled at identifying horses that have true learning deficits.

Sample Size

Achieving large sample sizes in equine research is challenging, with costs often a major limiting factor. This challenge is greater when both horses and riders are involved. Participants (human and equine) will often be recruited from the equestrian industry and it is unlikely that large groups of horses that must be fed, housed, and ridden homogeneously will be accessible. Small samples can be useful if all characteristics of the horses and riders involved are recorded fully (see *Sample characteristics and variables relating to rider and environment*). Depending on the study it may be necessary to standardize features of the horse such as age, breed or gender. Repeated measures designs reduce the effect of individual variation and allow horses to be used as their own controls. The potential

effect of different riders needs consideration within the design of the study. One rider will maximize consistency across horses but the results of the study might be attributable to the skills and experience of that individual rider. When multiple riders are used, the effect of different horse-human combinations must be considered. The controls to mitigate for rider effects will depend on the aim of the study and the resources available.

Where multiple treatments are involved, sample sizes should be increased unless treatment order effects can be controlled for by using a crossover design. Each treatment can leave a legacy and although wash-out periods between treatments help to reduce this effect, they do not guarantee that each horse is at a baseline condition at the start of the next treatment.

The optimal sample size balances statistical power with the practicality of accessing a large sample group (Eng, 2003). Unfortunately, *a priori* sample size calculation has been neglected rather commonly in equitation research where practical considerations largely determine the number of studied subjects. Researchers are strongly advised to conduct sample size calculations when they are planning a new experiment and to report the methods in their publications. Numerous guidelines exist for calculating power and/or sample size for various research designs (e.g. Cohen, 1988; Lenth, 2001). Additionally, free software is available to conduct power and sample size calculations for various research designs (e.g. Faul et al., 2007).

Standard deviations of the measures taken should be presented in the results to enable researchers to calculate the optimum sample size for future studies. Statistical analysis should be considered during planning and before data collection begins. In addition, the value of pilot studies cannot be over-emphasized. Researchers are strongly advised to seek expert statistical guidance in the final design of studies *prior* to starting data collection.

Randomization and blinding of trials

In any experiment variable levels of skill and bias in the experimenters and other personnel, such as riders and handlers, may affect the results (Kuehl, 2000; Morris, 2010). When interventions may influence data (e.g., training programs, physiotherapy), studies should be blinded to the extent possible until after data collection/analysis is completed. A further challenge is that when riders know what is being measured they may ride differently from normal, so riders should also be blinded. This may simply involve, for example, not knowing whether the pressure sensor is actually connected to the computer.

Sample characteristics and variables relating to rider and environment

Many variables may influence the behavior of the horse–human dyad. These variables should be clearly described, in a standardized way, in any study involving ridden-horse behavior and should include characteristics of horse, rider and environment. Table 1 provides a list of these variables and

a suggested mode for reporting them. The list is not exhaustive but includes the main factors to be reported. For example, skill level in riders is an important variable in equitation science studies.

By including details of these experimental variables in a consistent way, the comparison of different studies will be facilitated. With reference to *Ethical considerations and informed consent*, human participant details should not be attributable to individuals and must be stored securely. Compliance with relevant legislation relating to data protection is essential.

Table 1
Factors to be recorded and suggested mode of recording for each factor—Riding discipline, horse/pony (sample) characteristics, rider features, and environmental factors

Factor	Suggested mode of recording
Riding discipline	For example: English (Dressage/Show jumping/Eventing/Driving), Western (reining, pleasure, trail, horsemanship, cutting, and team penning), gaited horse riding, Baroque riding, working equitation, recreational riding
Horse/pony	
Age	Years (months)
Gender	Mare/gelding/stallion State whether the females are, or may be, pregnant
Breed/type	Specific breed or, at minimum, type: Hot-blood, Warm-blood or Cold-blood Where crossbreeds are being used, state the breeds involved only where there is documented evidence
Past experience	Time span undersaddle, previous training and related discipline, and highest competition level If applicable, state different use of horse before current use at the time of the experiment (e.g., was it formerly a trotter or a breeding mare?)
Current role	Competitive discipline (if this varies between horses/ponies and/or if they participate in more than one discipline, including level), perceived level of training and noncompetitive role
Management style	Type and amount of turnout, as well as companionship of other horses and seasonal variation Approximate size and type of turnout in SI units (i.e., pasture/paddock); hours/week of turnout. State whether turnout is in groups or solitary Type of housing: Individual or group housing, size of stall. If individual, description of possibilities for contact with other horses; if group, size of group
Type and amount of feed	Reported in kilograms per day for roughage, grains and supplements, and seasonal or temporal variation
Horse health	Horses should be clinically sound
Familiarity with test area	Circumstances (e.g., daily riding training) and time span (e.g., for 5 years)
Rider	
Gender	Male/female
Age	Years, range, mean age (\pm standard deviation)
Height/weight	SI units (meters/kilograms)
Riding experience	Time span (years or months), equestrian qualifications (specify where appropriate), range/discipline(s), competitive level, and success (provincial/local, national, or international). Professional role within equestrianism
Familiarity with horse	Time span and hours/week riding and interacting with horse/pony. Is the horse/pony cared for by the rider? Regularity of contact (daily, more/less frequent; e.g., perceived experience in riding). Involvement in training the horse (yes/no/partly)
Environmental factors	
Presence/absence of other horses	Group/individual test. Proximity of other horses/horse-rider combinations
Test area	Indoors/outdoors, riding surface, area; boundary type (open fencing, close-boarded etc.)
Environmental conditions	As appropriate for study: Ambient temperature (degrees Celsius); humidity; wind speed; light conditions
Extraneous variables	Noise, visual stimuli, any other unexpected event (description of event and frequency of occurrence)

Null findings should be reported in order to avoid repeats of studies that have already been carried out but have not been published. They will also inform future studies and can provide valuable insights into equitation *per se*.

Many of the above-mentioned variables will have more than two levels, so analyzing the influence of these variables and/or multiple treatment effects calls for consideration of multiple comparisons in the statistical analysis. Although some argue that it may not be necessary (Rothman, 1990; Perneger, 1998) to adjust for multiple comparisons, the best approach involves assessing each situation for risks of type I and type II errors and deciding whether adjustment for multiple comparisons is needed. The use or non-use of adjustment methods should be clearly stated and decisions to not use multiple comparison adjustments should be justified by the authors.

Behavioral and physiological data

Recording ridden-horse behavior

The evaluation and manipulation of the horse's behavior is fundamental to equitation and equestrian sports (Hall et al., 2013). The behavior of the horse is the final manifestation of all the processes active in its brain and body, their interactions and, in the ridden horse, is directly influenced by human action. To report on equine behavior and to discuss it, a clearly defined and widely understood and accepted definition of behavior units is required. At this point, such a basic ethogram of the domestic horse during interactions with humans is not available, although the equid ethogram by McDonnell (2003) is an important standard. Lehner (1996) defines an ethogram as a comprehensive set of descriptions of the characteristic behavior patterns of a species. Martin and Bateson (2007) emphasize that quantitative recording of behavior should be preceded by a period of informal observation, aimed at understanding and describing both the subjects and the behavior of interest. This first phase of observing and clearly describing behavior permits formulation of relevant research questions for quantitative and experimental research.

The treatise of Xenophon on *Horsemanship*, written more than 23 centuries ago, is widely referred to as one of the oldest documents on ridden horse behavior containing information that is still relevant today (Morgan, 1993). Currently the Fédération Equestre Internationale (FEI) produces guidelines for assessing ridden-horse behavior in the discipline of dressage, in which detailed descriptions of the movement involved and assessment criteria are provided (FEI, 2012a). The dressage movements described are all based on natural movements of the horse (Ödberg and Bouissou, 1999). However, within the FEI guidelines only desired behaviors are described in detail and there is limited reference to specific behavioral events. Although it is important to incorporate currently-used equestrian behavioral descriptors, a more comprehensive ridden-horse ethogram is required for use within experimental studies.

Table 2 provides an overview of studies from the area of equitation science that include an ethogram. For adaptation to the individual situation, each paper assessing ridden-horse behavior should, essentially, contain its own ethogram of ridden-horse behavior. However, using a common basis would be advantageous to compare different studies, and the list of definitions by McGreevy et al. (2005), in addition to descriptors included in the FEI Guidelines (FEI, 2012a), can provide a starting point for developing a standard ethogram for ridden-horse behavior.

Before behavior units or action patterns can be studied to determine their function, causation or ontogeny, they must first be clearly described. This starts with the description of the movement of body parts and combination of body parts leading to an assessment of the whole animal. Objective nomenclature generally demands the use of the Nomina Anatomica Veterinaria (ICVGAN, 2012), with a lay glossary made available for use by the entire equestrian community. Lehner (1996) states that since functional descriptions can be confusing and misleading, they should be avoided, except when the function is intuitively obvious or supported by data. This emphasizes that any functional

interpretation should always be supported by data on the relation with physiological, social or environmental parameters. A functional ethogram is necessary for many studies, but it needs regular updates to incorporate the rapidly evolving knowledge of functional interpretation of behavior units.

Table 2
Overview of studies providing an ethogram or descriptions of equine behavior during riding or handling and husbandry situations

References	Research question	Description of behavior	
		Description only	Illustrations included
Ridden work			
Egenvall et al. (2012)	Influence of training technique on horse behavior	++	-
Hall et al. (2013)	Review	-	-
Hall et al. (2014)	Relation between behavior, professional judgment and physiological measures	+++ / +++	-
Heleski et al. (2009)	Influence of equipment on conflict behavior	++	-
Kaiser et al. (2006)	Stress in therapeutic riding horses	++	+
Kienapfel (2011)	Influence of head/neck position on behavior	++	++
König von Borstel et al. (2011b)	Relation between behavior and personality traits	++	-
Lesimple et al. (2010)	Influence of rider action on equine welfare	++	+
Quick and Warren-Smith (2009)	Effect of bridles on equine behavior	+	-
von Borstel et al. (2009)	Influence of training technique on equine welfare	++	+
Handling situations			
Christensen et al. (2011)	Voluntary acceptance of rein tension	++	-
Innes and McBride (2008)	Comparison reinforcement techniques	++	-
Mansmann et al. (2011)	Behavior problems during farriery	+	-
McGreevy et al. (2009)	Review	+	-
Normando et al. (2011)	Factors influencing problem behavior	+	-
Visser et al. (2001)	Temperament assessment	++	-
General husbandry			
Arnold and Grassia (1982)	Agonistic interactions	++	-
Clegg et al. (2007)	Characterization of behavior in stereotyping horses	++	-
Garcia et al. (2010)	Influence of sampling frequency on behavior frequencies	++	-
Harewood and McGowan (2005)	Response of naïve horses to stabling	++	-
Heitor et al. (2006a)	Social dominance and functions of aggression	++	-
Heitor et al. (2006b)	Affiliative and sexual behavior	++	-
Heleski et al. (2002)	Influence of housing on weanling behavior	++	-
Jorgensen et al. (2009)	Social interactions in relation to gender composition of groups	++	-
Mal et al. (1991)	Response to confinement and social isolation	++	-
McDonnell and Poulin (2002)	Ethogram of play behavior	+++	-
McDonnell and Haviland (1995)	Agonistic interactions	+++	+
Peters et al. (2012)	Anticipatory behavior	+++	-
Strand et al. (2002)	Behavior during social isolation	++	-
Tateo et al. (2012)	Artificial rearing	++	-
Guidelines within equestrian sport			
Federation Equestre International (2012a)	Guidelines for evaluation of dressage performance	+	-
Federation Equestre International (2012b)	Rules for jumping events	+	-

(Key: - = absence of this type of information; + = availability of a small amount of descriptions or descriptions for just a few behavior patterns; ++ = availability of more detailed descriptions of several behavior patterns; +++ = provision of a detailed and exhaustive ethogram).

The methods to develop a functional ethogram should be sound, evidence-based and repeatable. Taylor and Mills (2006) described clearly which steps to take for assessing dog behavior. It is suggested that these steps are valid for all animal species and should be considered in every study.

For an ethogram or protocol to be reliable and valid, standardization of the assessment is a minimum requirement. To do this, all potential sources of variability must be identified and controlled or at least taken into account and reported. Standardization also relates to the formalization of reporting the animal's response by the use, for example, of check sheets, videos or voice recordings.

The reliability of the assessment demands that the researchers describe how well observers are trained to reliably and consistently score the behaviors in the ethogram. This is a fundamental for both inter- and intra-observer reliability. Differences between observations can be caused by imprecise definitions in the ethogram that allow too much room for interpretation. In all studies, reliability of the observers should be tested in a pilot study and stated. Usually, this is reported using Kappa or Kendall correlation coefficients. Repeatability, calculated on the basis of variance

components from mixed-model analyses, can also be used and has the advantage that the values are corrected for other potentially influencing factors. It is generally accepted that a Kappa or Kendall value of 0.41-0.60 is regarded as average concordance, a value of 0.61-0.80 as substantial and above 0.80 as excellent concordance (Burn et al., 2009).

Validation of the behavioral units should demonstrate how they reflect precisely what they are intended to measure. Correlating the results with physiological measures and/or other scoring techniques, such as questionnaires, and assessing repeatability of the behavioral elements across similar situations may provide additional support for the validity of a behavior pattern. For training purposes, video images may be essential, for example to test the reliability of observers over time. For video images to be useful for quantitative studies, one suggestion has been to maintain the horse's midline perpendicular to the lens. Cameras should be stationary (mounted on a tripod). Color footage with a minimum of 25 frames per second is preferred. Both the horse and the handler/rider should be completely visible at all times. In cases of long-reining or driving, this may require the use of two synchronized cameras. Multiple synchronized cameras can also record behavior from different angles simultaneously. To promote international uniformity and standardization of behavioral assessment, the International Society for Equitation Science (ISES) is collecting representative examples of behavior units to be included in an ethogram. It is anticipated that these examples will be made available online at the ISES website (www.equitationscience.com) and will be updated if better examples become available.

The descriptive definitions of behaviors can and should be a starting point for research on proximate and ultimate causes, functional attribution, physiological correlations or ontogeny. An ever-growing array of technical instruments to measure physical and biochemical variables allows detailed quantitative studies on the interpretation of the described behaviors. Techniques such as accelerometry, morphometry, kinematics, heart rate, infra-red thermography or assays of hormones, neurotransmitters or other biochemical markers, will help to validate and elucidate the functional interpretation of behavior units. This will lead to an equine ethogram with evidence-based functional definitions.

To facilitate comparability between studies, behavior patterns should be reported in frequencies (or duration, where appropriate) per hour of riding, or calculated as percentages of total riding time.

Physiological measures

For an accurate interpretation of behavior, the effect of training and equipment on the horse and for evidence based recommendations, objective measures are required. To evaluate the (emotional) response of the horse to ridden work, a number of physiological measures are commonly used, with the proviso that factors such as exercise level, fitness and age be taken into consideration when interpreting the results. Most importantly, the relationship between physical activity and changes in heart rate, heart rate variability (HRV) and hormonal responses must be assessed. Environmental factors, such as footing and climatic conditions, can influence the level of physical effort and needs

to be taken into account. Moreover, since it has been shown that horses demonstrate individual variation in heart rate parameters (Visser et al., 2002), good controls within the studies are necessary. Studies by Schmidt et al. (2010a) and Jansen et al. (2009) provide examples of how this measure has been used to monitor emotional responses in ridden horses. Other measures currently used are changes in heart rate variability (HRV), cortisol release and, more recently, changes in surface temperature as measured by infrared thermography (IRT).

A recent review of methods to validate methods for assessment of equine personality cites studies that measure both behavior and physiology in potentially stressful situations (*see* Table 2 in König von Borstel, 2013). In these studies correlations between physiological and behavioral data are in a medium to low range.

Heart rate variability (HRV): Little is known regarding thresholds at which stress, as assessed by the different HRV parameters, becomes negative (i.e., distress) and when it becomes pathological. So when some method results in higher levels of physiological stress, based on HRV measurements, it may be too early to conclude that that method is less desirable or more harmful.

As human mental states are known to have measurable effects on horse heart rate (e.g. von Borstel et al., 2005, König von Borstel, 2008), all participants and observers should ideally be blinded to treatments. Unless standards for error correction in heart-rate raw data are determined beforehand, the statistician/person dealing with the data should ideally also be blinded to treatments.

a. What endpoints to measure: It is important to ensure that endpoints are biologically, clinically and behaviorally meaningful. The Task Force of the European Society of Cardiology (1996) gives some recommendations regarding the minimum lengths of measurement periods for the different HRV parameters to be valid. This relates to human HRV, but since the underlying physiological mechanisms tend to be similar across species; these values may be a good starting point for horses, too.

Assessing repeatability of equine HR(V) data in various situations is currently topical (König von Borstel et al., 2011b; 2012). High repeatability *per se* does not demonstrate the validity of a method, or of using a certain time-frame to obtain a measurement, but low repeatability does demonstrate lack of validity.

With the low-frequency/high-frequency (LF/HF) measurement, there is also limited consensus on where to draw the lines for LF and HF bands. Software such as Kubios for analysis of HRV data uses the human range as the default setting, but they may not be appropriate. For example, Physick-Sheard et al. (2000) suggests frequency bands of (LF) >0.01– 0.07 and (HF) >0.07– 0.5 cycles/beat in horses.

b. The best current tools for measurement: The Holter monitor traces the electrocardiogram (ECG) continuously for a period of 24 hours or longer. It represents the gold standard, but it is expensive and can be impractical due to its limited mobility. With the fast improving technologies, it is expected

that these devices will improve considerably in the next few years. Polar™ equipment is, according to validating studies in humans (e.g., Macfarlane et al., 1989; Thivierge and Leger, 1989; Kingsley et al., 2005; Gamelin et al., 2006), among the most accurate available on the market. Significant problems with some Polar heart rate monitors have been reported (see Gorecka-Bruzda et al., 2012; Suwala et al., 2012), so it is suggested that each device should at least be cross-validated against a second device.

c. A standard way to calibrate and/or validate cardiac recordings: The standard way to validate the HRV-measurement tools is to compare measurements with ECG-recordings that have been obtained simultaneously. Thivierge and Leger (1989) considered correlation coefficients of 0.85 and higher and standard errors of the estimates and systematic errors of less than 5% of the average ECG values as requirements for a heart-rate monitor to be considered valid.

d. Outliers and missing values in HRV data: One of the main questions is how to deal with outliers both in the original data and in means and other measurements calculated from them. Some common correction methods include the use of some filtering settings provided by the manufacturer's software, or the removal of individual values outside the physiological range. However, there is no consensus on which methods are the most appropriate. Since the biologically relevant difference is not known for the different HRV measurements, it is difficult to deduce required sample sizes. However, perhaps the most important issue is that we still need to develop consensus on how to recognize and deal with artifact in the raw data. It would also be helpful to develop protocols on how to handle data and to control for multiple comparisons.

Missing data are fairly common due to technical failure and so are a cause for concern. Accordingly, appropriate statistical tests that can handle missing data should be chosen for analysis.

Some question whether HRV values should be corrected for mean heart rate. There is some evidence that the relationship between behavioral parameters and HRV parameters changed in a few cases when mean HR was accounted for in the mixed model, in contrast to correlations between the behavior patterns and HRV parameters that did not correct for mean HR (König von Borstel et al., 2011b). As HRV partially depends on HR, accounting for HR is judicious since it is not fixed.

Salivary cortisol: Elevation of cortisol is currently accepted as one measure of stress. Fecal cortisol metabolites can be used non-invasively to assess long-term stress (Schmidt et al., 2009), but more transient increases in cortisol release are measured in plasma and saliva. Unbound, biologically available plasma cortisol rapidly diffuses through the cells of the salivary glands and close correlations have been found between free cortisol in plasma and salivary cortisol concentrations in species including the horse (Peters et al., 2011). The time taken for plasma and salivary cortisol concentrations to increase post-stressor is comparable (Hughes and Creighton, 2007). It has been noted that baseline cortisol concentrations tend to increase with age (Hall et al., 2014) and training (Fazio et al., 2006), highlighting the need to take baseline measures against which changes can be assessed and to record all details of the horses concerned to enable valid conclusions to be drawn.

The duration of the stressor affects increases in cortisol concentration and the scales over which they should be measured (Schmidt et al., 2010a; Schmidt et al., 2010b).

Intervals for sampling saliva should be set at regular intervals throughout the study, with samples taken prior to set trials (e.g., a baseline) and at intervals during/post-trial. The cost of laboratory analyses may limit the frequency of sampling, but to enable comparison between studies, sampling should continue until 60 minutes post-study. Collection of saliva is generally by means of a cotton roll (e.g., salivettes, Sarstedt, UK, that have been adapted to prevent the horse from swallowing them) inserted in the mouth and that the horse is allowed to chew on. The saliva/salivettes are then frozen at -20°C until laboratory analysis using an enzyme-linked immunosorbent assay (ELISA). Generally inter- and intra-assay coefficients of variability are reported to demonstrate the reliability of the findings.

Infra-red thermography (IRT): Thermal imaging cameras produce visible images of radiation in the infrared range of the electromagnetic spectrum that is generally perceived as heat. Thermography has been used in medical/veterinary applications to monitor changes in body surface temperature and to detect early, pre-clinical signs of disease. Further discussion of previous applications of this technology, limitations and future potential can be found in the review by McCafferty (2007). Changes in skin temperature have also been shown to be associated with emotional/stress-related responses in humans and other animal species. There is currently considerable interest in the potential for IRT to provide a non-invasive, immediate means of objective assessment based on the premise that rapid changes in blood flow due to sympathetic activation affect the amount of radiated heat which can be measured as changes in surface temperature (Stewart et al., 2007).

If IRT is to provide an immediate, non-invasive measurement of stress in the ridden horse, then a number of factors must be taken into consideration.

a. Environmental factors: Eddy et al. (2001) recommended that a controlled environment (e.g., draught-free environment with low-level lighting and an ambient temperature of less than 30°C) is necessary to ensure the accuracy of thermographic scanning. In most applications involving ridden horses/field testing, including controlled ambient conditions is unrealistic; however, where possible, readings should be taken in a constant environment (e.g., within an indoor arena or barn) and not in direct/bright sunlight, draughty/windy conditions or high ambient temperatures. To determine the potential effect of ambient conditions on subsequent data, environmental temperature and humidity should be recorded at each IRT reading.

Environmental conditions at the time of IRT are crucial to obtain reliable measurements. Fluctuations of eye temperature within several seconds are commonplace and may be due to environmental factors (e.g., sun, wind, evaporation), presence of tears, cleanliness and the management of the horses prior to measurement. It is well known that practices such as rugging, cooling, complete rest, intake of food and water can affect core and eye temperature. These variables should be reported as

should the details of whether the horses were housed inside or otherwise at the time of recording the eye temperature.

b. Why eye temperature? Decreases in body surface temperature in response to anxiety/fear-provoking situations have been recorded in a number of species, including humans (finger temperature, Ekman et al., 1983), sheep (ear-pinna temperature, Lowe et al., 2005) and horses (lower limb, Eddy et al., 2001; ear-pinna temperature, Kay and Hall, 2009). Circulatory changes in these areas undoubtedly reflect emotional state but, when measured using IRT, a number of factors relating to the surface being assessed will affect the readings. It was noted by Eddy et al. (2001) that for thermographic images to be accurate and clear the horse should have a clean, dry coat that had not been groomed within 2 hours before the scan. Also, no topical agents should have been applied and exercise and sedation should be avoided because of their effects on peripheral blood flow (Eddy et al., 2001). Coat features such as color, thickness and moisture content can be sources of error (McCafferty, 2007).

Changes in temperature of the eye and surrounding area have been associated with aversive procedures. Since this area is less affected by coat variations, it offers the best potential for accurate monitoring of stress responses. Acute stress responses of cows during disbudding produce a magnitude and duration of initial drop in eye temperature following disbudding without anesthetic was comparable to that found following a fright in cattle (Stewart et al., 2008; Schaefer et al., 2006).

An initial drop in eye temperature has been recorded in one horse following a visual startle test (Yarnell, 2012). See Figure 1. The change in temperature of the left eye was recorded over a period of 5 seconds (–1s from presentation of umbrella to +4s post presentation). The speed and duration of the response indicates that subtle temperature changes in response to stressors may be missed. The magnitude of change was greater than that found in cattle in response to pain (Stewart et al., 2008). Movement away from the thermal camera may have contributed to the change, further emphasizing the need for controlled conditions. The duration of the response appears associated with the duration of the stressor.

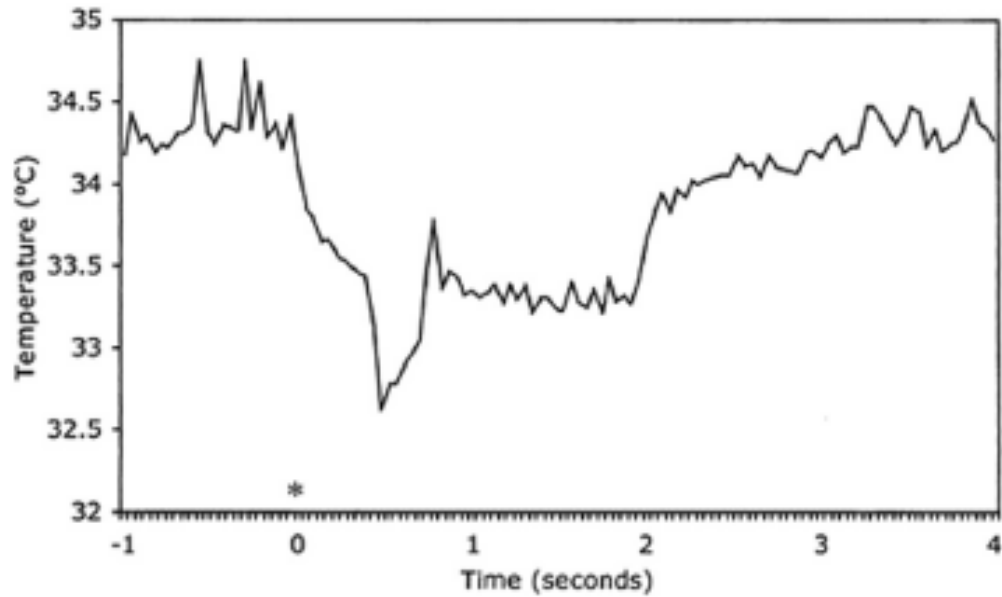


Figure 1. Change in eye temperature of one horse following visual startle test (opening of umbrella) at time 0 (indicated by *), recorded using a FLIR Thermovision A40M thermal camera (sensitivity: 0.08°C at 30°C, range: -40°C-500°C, image capture: 30 frames/second; Yarnell, 2012).

c. Recording eye temperature: Reliable and repeatable eye temperatures can be recorded only if taken from a set distance of between one and two meters and at an angle of 90°. It is also useful to provide the following details of the temperatures (°C): range, maximum, mean maximum and standard deviation. Temperatures should be collected from a specified area. Valera et al. (2012) provide images showing the location and value of the maximum temperature recorded, and concurrent atmospheric temperature and relative humidity data. The lachrymal caruncle, the region associated with maximum temperature, has also been associated with responses to pain and stress (Valera et al., 2012). This is likely the optimum location for monitoring affective state in the horse, though validation is required. The maximum temperature of each eye should be recorded within this area and if no significant difference is found between right and left eye, mean maximum eye temperature can be used.

d. Comparison with other measures: Eye temperature has the potential for providing an objective, non-invasive measure of “stress” but still requires validation with currently accepted measures (heart rate, HRV, salivary cortisol). Correlations between maximum eye temperature and salivary cortisol responses in horses have been found (Cook et al., 2001; Yarnell et al., 2013), although some variation is apparent (Valera et al., 2012). More detailed comparisons of the timing, duration and extent of the temperature change and the autonomic (sympathetic) and hypothalamic-pituitary-adrenal responses are required to fully interpret IRT data.

One other notable correlation that requires further investigation is that found between eye temperature and core body temperature. Stress-induced increases in core body temperature in humans and other animals (Oka et al., 2001) range from 0.6-2°C, depending on stressor and species (Oka et al., 2001). Thermographic eye temperature has been associated with core body temperature in ponies, so IRT has been suggested as a quick and non-invasive method of predicting fever onset (Johnson et al., 2011). Comparable increases of both eye (1.5°C) and core body temperature (0.35°C) have been found when horses are lunged in a training aid (Pessoa) following the tightening of the equipment (Hall et al., 2011). Physiological and physical responses to stress appear to result in body temperature changes, so core body temperature should be monitored where possible.

Physiological data collection may facilitate accurate interpretation of behavior. Maximum eye temperature has been shown to correlate with handling scores, with higher eye temperatures being associated with more anxious behavior (Hall et al., 2014). Some IRT equipment can collect continuous images (video option e.g., MOBIR® M8 thermal camera used in Hall et al., 2014). When still and video images were compared, eye temperatures from static images at a controlled distance were significantly higher than those obtained from the video footage. The distance from the horse, associated movement and variable angle affect temperatures recorded from video. Camera readings are impossible in bright sunlight (static or video), so refinements are needed for field condition measurements.

e. Validation of equipment: Validation of equipment used is important. Few studies report how readings from the equipment compare with alternative measures of surface temperature, but results affect whether comparisons can be made. The data provided by the range of equipment now available are unlikely to be consistent/comparable between models or conditions. For example, even the length of time the camera has been switched on can affect the readings. Repeated validation trials throughout the study should confirm the reliability of readings taken.

The specification of the equipment used must be included in all studies, with reference to sensitivity, range and image-capture rate. The use of IRT to monitor body surface temperature has certain advantages over other methods, such as temperature loggers: temperatures obtained via sensors attached to the surface being monitored may be compromised by the insulating properties of the equipment; non-contact/laser thermometers record the temperature at a single point only. High specification equipment is costly but necessary for body surface temperature changes to be used as a means of assessing stress in the horse. The temperature changes found in other species (associated with emotional responses in particular) have been small and rapid and would not be recorded by many thermal cameras. The advantage of immediate data, which is easy to collect, must not be overshadowed by a lack of accuracy. IRT has real potential in the assessment of stress but a great deal of work is needed before the results of this approach can be relied upon and applied more widely.

Measurement of human-horse interaction

Measuring interactions within a dyad means that the responses of both contributors must be assessed at the same time. Such measures are challenging because biofeedback from the horse may affect the rider's behavior and vice versa. Numerous innate and acquired variables in both members of the dyad affect the results. Furthermore, assessment of equestrian technique can be based on undefined terms like "feel". Early examples of measurements include rein tension (Clayton et al., 2005; Manfredi et al., 2005; Warren-Smith et al., 2007; Heleski et al., 2009; Manfredi et al., 2009; Clayton et al., 2011; Kuhnke et al., 2010; Christensen et al., 2011; Hawson et al., 2014, König von Borstel and Glißmann, 2014), leg pressure (Nevison et al., 2011, Hawson et al., 2012) and pressure from the seat (de Cocq et al., 2010b; Clayton et al., 2013). These studies were conducted to take the guesswork out of equitation, to identify what works and what does not and to ensure that, wherever possible, techniques that compromise horse welfare are addressed. Ultimately, this will help us define best practice. The emergence of smart textiles will help us study the human-horse interface (McGreevy et al., 2014), and together with the cooperation of some of the world's best performing competition riders, it is likely that what cannot yet be measured will be revealed.

Rein-tension measurements

Currently available equipment

To date, three rein-tension devices have been used in peer-reviewed research. These are the Mini Low Profile (MLP)-75™ load cell (Transducer Technologies, Temecula, CA, USA) (Clayton et al., 2011; Heleski et al., 2009; Manfredi et al., 2009), the MLP-100™ (Transducer Technologies, Temecula, CA, USA) (Clayton et al., 2003; Manfredi et al., 2005) and ReinCheck™ (Crafted Technology, Sydney, NSW, Australia), derived using Signal Scribe™ (Crafted Technology, Sydney, NSW, Australia) (Edwards and Randle, 2010; Kuhnke et al., 2010, Christensen et al., 2011; Randle et al., 2011;).

The accuracy of what can be measured is still being debated. Recording rein tension for analysis and display requires that the instrumentation converts the analogue tension into a digital form, a process that results in inherent loss of data. Display of the values is limited by the number of digital bits available to form the digital values representing the analogue-based tension (Staller, 2005). Increasing the number of bits results in greater resolution and precision. Quantization error is also inherent in analogue-to-digital conversions of time-variant signals (Staller, 2005). The effect of such errors on our current understanding of rein-mediated interactions is unclear, requiring ongoing recording of limitations of instrumentation.

Range and sampling rate

It is important that all tension meters are capable of measuring across the full range of tensions encountered in the reins of the ridden horse. The range capabilities of tension meters varied within studies, as did the range of tensions reported in the ridden horses. The highest recorded tension was

104 N by Clayton et al. (2005). Horses will voluntarily tolerate tensions of up to 11 N (Christensen, 2011), suggesting that tensions greater than this cause discomfort.

Sampling rate is an important factor in measuring rein tension. The physical rein tension is presumably changing continually while the digital signal is determined by how often the signal is recorded. If the sampling rate is too slow, critical changes in rein tension may not be observed. Clayton et al. (2011) confirmed Clayton et al.'s (2005) earlier report that the horse's stride produced a regular pattern of peaks in rein tension. In Starke et al.'s (2012) study, the horse's shortest mean stride duration of 683 ms was when trotting in a straight line. This provides a stride frequency of approximately 1.5 Hz (cycles per second) or 90 beats per minute. In contrast, stride frequencies reported by Clayton were 55 strides per minute for walk (Clayton, 1995), 79 strides per minute for trot (Clayton, 1994a), and 99 strides per minute for canter (Clayton, 1994b). There are two peaks in rein tension for every stride cycle in walk and trot (frequency 1.8Hz and 2.6Hz respectively) and one per stride cycle in canter (frequency 1.7 Hz) (Clayton et al., 2005).

Both horse and rider can also change the tension of the reins through muscular contraction. Desmedt and Godaux (1977) demonstrated that, in humans, motor unit discharge rates can be up to 120 Hz during contractions associated with throwing objects but that these may be less than 30 Hz in slow contractions. Horses are capable of fast movement, so it is likely that equine motor unit firing rates and consequent muscle movement can occur at a fast rate. Tension meters should sample rein tension at a similar rate as muscle contractions by both horse and rider can change rein tension.

Validation and Calibration

Humidity and temperature may affect the accuracy of the strain gauges within rein-tension meters. These are exactly the types of conditions that will change as horses warm up during ridden work. Validation of all rein-tension meters under practical conditions is necessary to ensure that the instruments are measuring within the agreed error range.

Calibration by suspending a known range of weights from the reins was first described by Clayton et al., (2003) and has been used in many studies. The range of weights should reflect the range of tensions expected to be recorded, currently approximately 0-11kg. It is unclear if conducting calibration in the vertical plane of instruments that will measure tension in a more horizontal plane affects accuracy. Researchers may need to consider alternative calibration techniques. Accurate and adaptable calibration systems must be incorporated into the design of rein-tension devices.

Until we can verify calibration techniques and be confident that the gauges are not adversely affected by changes in humidity and temperature, we have to recognize that, at best, we are reporting approximations.

Wireless

Making rein-tension meters wireless will minimize the equipment on the horse and provide real-time feedback to researchers, coaches and riders (McGreevy et al., 2014 and Warren-Smith et al., 2007). There should also be a capacity to link with real-time surveillance (video) so that data can be matched to observed events. This would make it easier to provide focused feedback to riders. It is important that this linked facility provides for external control of data collection by turning the logging on and off and permitting placement of event markers on the recording.

Limitations and future requirements

It is hoped that technology in this area will continue to develop. Manufacturers should consider the current recommendations relating to range, sampling rate, ease and accuracy of validation and calibration. Wireless communication and control are mandatory as is minimizing weight. Perhaps smart textiles development may eventually provide for tension measurement within the material of the reins themselves. It is worth noting that rein tension *per se* does not relate directly to pressure in the horse's mouth, and providing real-time biofeedback to the rider and/or coach remains elusive. Since bit design has a major impact on the conversion of rein tension to the intensity and distribution of pressures in the horse's mouth, it is imperative that the type of bit used is reported.

Pressure-sensing instruments for use under rider and/or saddle

Pressure-measuring devices designed to objectively determine saddle fit have been available for some time (Harman, 1994; Jeffcott et al., 1999). Pressure-sensing technology has been adopted from medical and mechanical applications. Several authors have garnered valid information for horse equipment design and fit (Baltacis et al., 2006; Byström et al., 2010; Fruehwirth et al., 2004; Hofmann et al., 2006; Kotschwar et al., 2010; Latif et al., 2010; Meschan et al., 2007; Monkemoller et al., 2005; Nyikos et al., 2005; von Peinen et al., 2010; Werner et al., 2002; Winkelmayr et al., 2006). More recently, attempts have been made to use this technology to measure changes in pressure on the horse's back from the rider's seat and legs (Belock et al., 2012; Byström et al., 2009; Byström et al., 2010; Clayton et al., 2013; de Cocq et al., 2008; de Cocq et al., 2009; de Cocq et al., 2010a, de Cocq et al., 2010b, Geutjens et al., 2008; Nevison et al., 2011; Nevison and Timmis, 2013; Peham, 2008; Peham et al., 2004; von Peinen et al., 2009).

Currently used pressure sensors

There are currently two types of pressure sensor used in saddle pressure pads. Several manufacturers supply systems based on piezo-resistive sensors. These sensors detect changes in the resistance that occurs when the sensor is deformed. Pullin et al. (1996) and de Cocq et al. (2006) have raised concerns about the accuracy of these sensors when used to measure pressures on horses' backs. Jansson et al. (2012) reported that these types of sensors suffer a reduction in load output when exposed to increasing humidity, such as is found under the saddle of the ridden horse.

The most popular system adopted by equine researchers uses a capacitive sensor. When a force changes the relative distance between two closely associated membranes, there is a corresponding change in the capacitance (Puers, 1993), which is converted to an electrical signal that is recorded. These sensors have the advantage of requiring low power, being more sensitive and less prone to drift or hysteresis than the piezo-resistive sensors (Puers, 1993). Unfortunately, these sensors, at least in saddle pressure pad format, are also prohibitively expensive.

Pressure-measurement challenges

Regardless of sensor type employed, the area between the ridden horse's back and the rider's seat is complex with several interacting curved and deformable surfaces. This results in multi-directional forces moving at different velocities defined by the horse's movement, the rider's weight and the rider's capacity to synchronize with the horse's movement (Lagarde et al., 2005). There are usually several layers of materials of different densities added to this already crowded space when a saddle is used. These include textile layer(s) of the seat of the saddle, structural materials associated with the frame of the saddle, shock-absorbing material and an external cover that contacts the horse's back. Commonly, riders use at least one further layer of external material between the saddle and the horse's back (Hawson et al., 2013). All these layers create shearing, rotational and normal forces at each surface interface. Current sensors can only detect forces/pressures acting perpendicularly to the sensor. The surface of the horse's back is curvilinear and moves independently of the sensor mat. It is likely that data on saddle or seat pressure to date have not captured the tangential forces/pressures associated with shear, and so underestimate the total force acting on the horse's back. To date, most researchers have reported ranges, peaks and means of pressures under the saddle in terms of total area under the saddle as well as sections of the saddle. Although this research has produced useful measures for saddle fit, it is apparent that more sophisticated analytical systems must be applied in order to identify changes in pressure on the horse's back associated with movement, and ultimately to decipher cues riders may be supplying through bodyweight redistribution.

The pressure-sensing technologies employed in saddle pressure pads are similar to those used in many smart phones and touch screens. It is likely that consumer-driven demand for better technologies will also give rise to improvements in saddle pressure pad products. Ideally these measuring devices will be wireless and provide real-time data on screen. The sensor pads must be thin and deformable to match a variety of horse shapes, yet sufficiently robust to survive the rigors of being between riders, saddles and horses' backs. The sensors themselves must be accurate and reliable in this environment. There should be no cross talk between each sensor. Calibration should be automated and fast. Data analysis must become more sophisticated and dynamic in order to capture the complexities of the interaction between horse, saddle and rider.

Calibration and validation protocols

There are two issues here. First, how each type of sensor is calibrated and how calibrations and other quality-control standards are reported. Second, how the different types of sensor compare with each other under standardized conditions, so the data sets of different researchers can be appropriately contrasted in meta-analyses.

What endpoints to measure

Appropriate SI units should be used. It is anticipated that future workshops will develop protocols that address sampling rates, cross-referencing with other equipment running concurrently, how much data can be extracted and how best to handle such data as data collation may cause a loss of detail and potentially important information.

Conclusion

Equitation science will be enhanced by consistent and robust methodology. Within equestrianism there is a long tradition of performance assessment, based on behavioral responses and assumptions regarding how this behavior should be interpreted. More recently, research has provided evidence that some of these assumptions may be flawed and that further investigation is required. The framework paper presented here is intended to stimulate the development and use of standardized methods to assess horses' responses in equitation science studies, in order to improve equine health and welfare. These guidelines are expected to be especially helpful for early career researchers. They may also allow more experienced scientists to make informed decisions on methodological issues. The findings can then be used to inform the future direction of equestrianism. As with any new discipline, the methods used are at an early stage of development and it is imperative that we develop methods to ensure that subsequent studies are valid and comparable. It is also imperative that we collaborate with equestrian practitioners to ensure that the findings will be utilized. We believe that the recommendations provide an important step in underpinning equitation science research and that the potential impact of future research will be enhanced. It is anticipated that this framework will be updated regularly as new technology and software emerge.

Acknowledgments

We thank Associate Professor Michal Schneider and Ms Lynn Cole for their help with earlier versions of this manuscript. Professor Paul McGreevy is thanked for his financial support of the inaugural ISES workshop.

Conflict of Interest Statement

No conflict of interest has been declared by any of the authors.

Authorship statement

All authors have materially participated in the work and have approved the final article. The idea for the paper was conceived by Paul McGreevy. The workshop on which this article is based was designed by Paul McGreevy, Alison Averis, Francis Burton and Kathalijne Visser. The project was undertaken by all authors. The paper was written by all authors. Besides the 4 equal first authors and the last, the others were arranged in alphabetical order. The project was entirely non-invasive and was funded by the International Society for Equitation Science and the workshop's attendees with additional funding from Paul McGreevy.

References

- Arnold, W., Grassia, A., 1982. Ethogram of agonistic behavior for thoroughbred horses. *Appl. Anim. Ethol.* 8, 5-25.
- Baltacis, A., Hofmann, A., Schobesberger, H., Peham, C., 2006. Evaluation of pressure distribution under a fitting saddle with different saddle pads. *J. Biomech.* 39, S559-S559.
- Belock, B., Kaiser, L.J., Lavagnino, M., Clayton, H.M., 2012. Comparison of pressure distribution under a conventional saddle and a treeless saddle at sitting trot. *Vet. J.* 193, 87-91.
- Burn, C.C., Pritchard, J.C., Whay, H.R., 2009. Observer reliability for working equine welfare assessment: Problems with high prevalences of certain results. *Anim. Welfare* 18, 177-187.
- Byström, A., Rhodin, M., von Peinen, K., Weishaupt, M.A., Roepstorff, L., 2009. Basic kinematics of the saddle and rider in high-level dressage horses trotting on a treadmill. *Equine Vet. J.* 41, 280-284.

- Byström, A., Stalfelt, A., Egenvall, A., Von Peinen, K., Morgan, K., Roepstorff, L., 2010. Influence of girth strap placement and panel flocking material on the saddle pressure pattern during riding of horses. *Equine Vet. J.* 42, 502-509.
- 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, 2, 223-238.
- Clayton, H.M., 1994a. Comparison of the stride kinematics of the collected, working, medium and extended trot in horses. *Equine Vet. J.* 26, 230-234.
- Clayton, H.M., 1994b. Comparison of the collected, working, medium and extended canters. *Equine Vet. J. Suppl.* 17, 16-19.
- Clayton, H.M., 1995. Comparison of the stride kinematics of the collected, medium and extended walk in horses. *Am. J. Vet. Res.* 56, 849-852.
- Clayton, H. M., Singleton, W.H., Lanovaz, J. L. Cloud, G. L., 2003. Measurement of rein tension during horseback riding using strain gage transducers: *Exp. Techniques* 27, 34-36.
- Clayton, H.M., Singleton, W.H., Lanovaz, J.L., Cloud, G.L., 2005. Strain gauge measurement of rein tension during riding: a pilot study. *Equine Comp. Exerc. Physiol.* 2, 3, 203-205.
- 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.
- Clayton, H.M., Belock, B., Lavagnino, M., Kaiser, L.J., 2013. Forces and pressures on the horse's back during bareback riding. *Vet. J.* 195, 1, 48-52.
- Clegg, H.A., Buckley, P., Friend, M.A., McGreevy, P.D., 2007. The ethological and physiological characteristics of cribbing and weaving horses. *Appl. Anim. Behav. Sci.* 109, 68-76.
- Cook, N.J., Schaefer, A.L., Warren, L., 2001. Adrenocortical and metabolic responses to ACTH injection in horses: An assessment by salivary cortisol and infrared thermography of the eye. *Can. J. Anim. Sci. (abstract of technical paper)* 81, 621.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.
- De Cocq, P., Muller, M., van Leeuwen, J., 2008. Horse-rider interaction: A simple model for different riding techniques at trot. *Comp. Biochem. and Phys. A* 150, S81-S82.
- De Cocq P., van Weeren P.R., Back W., 2006. Saddle pressure measuring: Validity, reliability and power to discriminate between different saddle-fits. *Vet. J.* 172,265-273.

De Cocq, P., Clayton, H.M., Terada, K., Muller, M., van Leeuwen, J.L., 2009. Usability of normal force distribution measurements to evaluate asymmetrical loading of the back of the horse and different rider positions on a standing horse. *Vet. J.* 181, 266-273.

De Cocq, P., Duncker, A.M., Clayton, H.M., Bobbert, M.F., Muller, M., van Leeuwen, J.L., 2010a. Vertical forces on the horse's back in sitting and rising trot. *J. Biomech.* 43, 327-631.

De Cocq, P., Mooren, M., Dortmans, A., Van Weeren, P.R., Timmerman, M., Muller, M., Van Leeuwen, J.L., 2010b. Saddle and leg forces during lateral movements in dressage. *Equine Vet. J.* 42, 644-649.

Desmedt, J.E., Godaux, E., 1977. Ballistic contractions in man: characteristic recruitment pattern of single motor units of the tibialis anterior muscle. *J. Physiol.* 264, 673-693.

Eddy, A.L., Van Hoogmoed, L.M., Snyder, J.R., 2001. The role of thermography in the management of equine lameness. *Vet. J.* 162, 172-181.

Edwards, H., Randle, H. 2010. Perceived rider rein contact and the effect on the horses' stride and step length at walk and trot (Abstract only). In: 6th Annual Equitation Science Conference, Hartmann, E., Zetterqvist, M., Fransson, C., Dalin, G., (Eds). Uppsala, Sweden.

Egenvall, A., Eisersiö, M., Roepstorff, L., 2012. Pilot study of behavior responses in young riding horses using 2 methods of making transitions from trot to walk. *J. Vet. Behav.* 7, 157-168.

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

Eng, J., 2003. Sample Size Estimation: How Many Individuals Should Be Studied? *Radiology* 227, 2, 309-313.

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behav. Res. Methods* 39, 175-191

Fazio, E., Calabrò, G., Medica, P., Messineo, C., Ferlazzo, A., 2006. Serum cortisol levels of Quarter horses: Circadian variations and effects of training and western riding events. In: Lindner, A. (Ed.) *Management of lameness causes in sport horses. Conference on Equine Sports Medicine and Science*, Cambridge, UK. Wageningen Academic Publishers, 175-179.

Federation Equestre International (FEI) (Ed.), 2012a. Rules for dressage events. 24th edition. [Online] <http://www.fei.org/disciplines/dressage/rules> [date accessed 21.3.2013].

- Federation Equestre Internationale (FEI) (Ed.), 2012b. Jumping rules. 24th edition. [Online] <http://www.fei.org/disciplines/jumping/rules> [date accessed 21.3.2013].
- Fruehwirth, B., Peham, C., Scheidl, M., Schobesberger, H., 2004. Evaluation of pressure distribution under an English saddle at walk, trot and canter. *Equine Vet. J.* 36, 754-757.
- Gamelin, F.X., Berthoin, S., Bosquet, L., 2006. Validity of the Polar S810 Heart Rate Monitor to Measure R-R Intervals at Rest. *Med. Sci. Sport. Exer.* 38, 887-893.
- Garcia, H.A.C., Furtado, C.E., Soncin, M.R.S.P., Daniel, F.W., Wandembruck, K.T., Polizel, V.P., Torrecilhas, J.A., 2010. Interference of withdrawal of notice ethogram for determining the behavior of subject to start of potros haltes and stabling. *Rev. Agrarian* 3, 162-168.
- Geutjens, C.A., Clayton, H.M., Kaiser, L.J., 2008. Forces and pressures beneath the saddle during mounting from the ground and from a raised mounting platform. *Vet. J.* 175, 332-337.
- Gorecka-Bruzda, A., Suwala, M., Jezierski, T., 2012. Typical artefacts in equine heart rate. In: Randle, H., Waran, N. Williams, J., (Eds.), 8th International Equitation Science Symposium 2012, Edinburgh, United Kingdom, 90.
- Hall, C., Burton, K., Maycock, E. and Wragg, E., 2011. A preliminary study into the use of infrared thermography as a means of assessing the horse's response to different training methods. *J. Vet. Behav.* 6, 291-292.
- Hall, C., Nuws, N., White, C., Taylor, L., Owen, H., McGreevy, P., 2013. Assessment of ridden horse behaviour. *J. Vet. Behav.* 8, 74-81.
- Hall, C., Kay, R., Yarnell, K., 2014. Assessing ridden horse behaviour: Professional judgment and physiological measures. *J. Vet. Behav.* 9, 22-29.
- Harewood, E.J., McGowan, C.M., 2005. Behavioural and physiological responses to stabling in naïve horses. *J. Equine Vet. Sci.* 25, 164-170.
- Harman, J.C., 1994. Practical use of a computerized saddle pressure-measuring device to determine the effects of saddle pads on the horse's back. *J. Equine Vet. Sci.* 14, 606-611.
- Hawson, L.A., 2012. Rider inputs. In: Randle, H., Waran, N. Williams, J., (Eds.), 8th International Equitation Science Symposium 2012, Edinburgh, United Kingdom, 90.
- Hawson, L.A., McLean, A.N., McGreevy P.D., 2014. Riders' application of rein tension for walk-to-halt transitions on an artificial horse. *J. Vet. Behav.* 9, 164-168.
- Hawson, L.A., McLean, A.N., McGreevy, P.D., 2013. A retrospective survey of riders' opinions of the use of saddle pads in horses. *J. Vet. Behav.* 8, 74-81.

- Heitor, F., Oom, M.M., Vicente, L., 2006a. Social relationships in a herd of Sorraia horses, Part I. Correlates of social dominance and contexts of aggression. *Behav. Process.* 73, 170-177.
- Heitor, F., Oom, M.M., Vicente, L., 2006b. Social relationships in a herd of Sorraia horses. Part II. Factors affecting affiliative relationships and sexual behaviors. *Behav. Process.* 73, 231-239.
- Heleski, C.R., Shelle, A.C., Nielsen, B.D., Zanella, A.J., 2002. Influence of housing on weanling horse behavior and subsequent welfare. *Appl. Anim. Behav. Sci.* 78, 291-302.
- Heleski, C.R., McGreevy, P.D., Kaiser, L.J., Lavagnino, M., Tans, E., Bello, N., Clayton, H.M., 2009. Effects on behavior and rein tension on horses ridden with or without martingales and rein inserts. *Vet. J.* 181, 56-62.
- Hofmann, A., Baltacis, A., Schobesberger, H., Peham, C., 2006. Evaluation of pressure distribution under a too-wide saddle with different saddle pads. *J. Biomech.* 39, S559-S560.
- Hughes, T.J., Creighton, E., 2007. Measuring cortisol as an indicator of stress response in domestic horses. *J. Equine Stud.* 4, 26-28.
- Innes, L., McBride, S., 2008. Negative versus positive reinforcement: An evaluation of training strategies for rehabilitated horses. *Appl. Anim. Behav. Sci.* 112, 357-368.
- Jansen, F., Van der Krogt, J., Van Loon, K., Avezzū, V., Guarino, M., Quanten, S., Berckmans, D., 2009. Online detection of an emotional response of a horse during physical activity. *Vet. J.* 181, 38-42.
- Jansson, K.S., Michalski, M.P., Smith, S.D., LaPrade, R.F., Wijdicks, C.A., 2012. Tekscan pressure sensor output changes in the presence of liquid exposure. *J. Biomech.* 46, 612-4.
- Jeffcott, L.B., Holmes, M.A., Townsend, H.G.G., 1999. Validity of Saddle Pressure Measurements Using Force-sensing Array Technology – Preliminary Studies. *Vet. J.* 158, 113-119.
- Johnson, S.R., Rao, S., Hussey, S.B., Morley, P.S., Traub-Dargatz, J.L., 2011. Thermographic eye temperature as an index to body temperature in ponies. *J. Equine Vet. Sci.* 31, 63-66.
- Jorgensen, G.H.M., Borsheim, L., Mejdell, C.M., Sondergaard, E., Boe, K.E., 2009. Grouping horses according to gender – effects on aggression, spacing and injuries. *Appl. Anim. Behav. Sci.* 120, 94-99.
- Kaiser, L., Heleski, C.R., Siegford, J., Smith, K.A., 2006. Stress-related behaviors among horses used in a therapeutic riding program. *J. Am. Vet. Med. Assoc.* 228, 39-45.
- Kay, R., Hall, C.A., 2009. The use of a mirror reduces isolation stress in horses being transported by trailer. *Appl. Anim. Behav. Sci.* 116, 237-243.

Kienapfel, K., 2011. And what are the opinions of the horses? On the expressive behavior of horses in different neck position. In German with English abstract. *Pferdeheilkunde* 27, 372-380.

Kingsley, M., Lewis, M.J., Marson, R.E., 2005. Comparison of Polar 810s and an ambulatory ECG system for RR interval measurement during progressive exercise. *Int. J. S. Med.* 26, 39-44.

König von Borstel, U.U., 2008. *Fear in Horses and Riders: Two Hearts Living in just One Mind – The Influence of Rider, Training and Genetics on horses' fear.* Verlag Dr. Müller, Saarbrücken, Germany.

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

König von Borstel, U., Euent, S., Graf, P., König, S., Gauly, M., 2011b. Equine behavior and heart rate in temperament tests with or without rider or handler. *Physiol. Behav.* 104, 454-463.

König v. Borstel, U., Peinemann, V., Glißmann, C., Euent, S., 2012. Willing to work? Suitability of different riding situations in evaluation of equine personality In: Randle, H., Waran, N. Williams, J., (Eds.), 8th International Equitation Science Symposium 2012, Edinburgh, United Kingdom, 90.

König von Borstel, U., 2013. Assessing and influencing personality for improvement of animal welfare: a review of equine studies. *CAB Reviews: Perspectives in Agriculture, Veterinary Science, Nutrition and Natural Resources* 8, 1-27.

König von Borstel, U., Glißmann, C. 2014. Alternatives to conventional evaluation of rideability in horse performance tests: suitability of rein tension and behavioural parameters. *PLoS ONE* 9: e87285

Kotschwar, A.B., Baltacis, A., Peham, C., 2010. The effects of different saddle pads on forces and pressure distribution beneath a fitting saddle. *Equine Vet. J.* 42, 114-118.

Kuehl, R.O., 2000. *Design of Experiments: Statistical Principles of Research Design and Analysis*, 2nd edition. Duxbury Press

Kuhnke, S., Dumbell, L., Gauly, M., Johnson, J.L., McDonald, K., von Borstel, U.K., 2010. A comparison of rein tension of the rider's dominant and non-dominant hand and the influence of the horse's laterality. *Comp. Exerc. Physiol.* 7, 2,57-63.

Lagarde, J., Kelso, J.A.S., Peham, C., Licka, T., 2005. Coordination dynamics of the horse-rider system. *J. Motor Behav.* 37, 418-424.

- Latif, S.N., von Peinen, K., Wiestner, T., Bitschnau, C., Renk, B., Weishaupt, M.A., 2010. Saddle pressure patterns of three different training saddles (normal tree, flexible tree, treeless) in Thoroughbred racehorses at trot and gallop. *Equine Vet. J.* 42, 630-636.
- Lehner, P.N., 1996. *Handbook of ethological methods*. 2nd edition. Cambridge University Press, Cambridge UK.
- Lenth, R. V., (2001). Some practical guidelines for effective sample size determination. *Am. Stat.* 55, 187-193.
- Lesimple, C., Fureix, C., Menguy, H., Hausberger, M., 2010. Human direct actions may alter animal welfare, a study on horses (*Equus caballus*). *PLoS one* 5, e10257.
- Lowe, T.E., Cook, C.J., Ingram, J.R., Harris, P.J., 2005. Changes in ear-pinna temperature as a useful measure of stress in sheep (*Ovis aries*). *Anim. Welfare* 14, 35-42.
- Macfarlane, D.J., Fogarty, B.A., Hopkins, W.G., 1989. The accuracy and variability of commercially available heart rate monitors. *NZ J. Sports Med.* 17, 51-53.
- Mal, M.E., Friend, T.H., Lay, D.C., Vogelsang, S.G., Jenkins, O.C., 1991. Behavioral responses of mares to short-term confinement and social isolation. *Appl. Anim. Behav. Sci.* 31, 13-24.
- Manfredi, J., Clayton, H.M., Rosenstein, D., 2005. Radiographic study of bit position within the horse's oral cavity. *Equine Comp. Exerc. Physiol.* 2, 3, 195-201.
- Manfredi, J., Rosenstein, D., Lanovaz, J.L., Nauwelaerts, V., Clayton, H.M., 2009. Fluoroscopic study of oral behaviors in response to the presence of a bit and the effects of rein tension. *Comp. Exerc. Physiol.* 6, 4, 143-148.
- Mansmann, R.A., Currie, M.C., Correa, M.T., Sherman, B., vom Orde, K., 2011. Equine Behavior Problems – Around Farriery: Foot Pain in 11 Horses. *J. Equine Vet. Sci.* 31, 44-48.
- Martin, P., Bateson, P., 2007. *Measuring Behaviour: An Introductory Guide*. 3rd edition. Cambridge University Press, Cambridge UK.
- McCafferty, D.J., 2007. The value of infrared thermography for research on mammals: previous applications and future directions. *Mammal Rev.* 37, 207-223.
- McDonnell, S.M., 2003. *The equid ethogram. A practical field guide to horse behavior*. The Blood Horse Inc; Lexington, Kentucky USA.
- McDonnell, S.M., Haviland, J.C.S., 1995. Agonistic ethogram of the equid bachelor band. *Appl. Anim. Behav. Sci.* 43, 147-188.

- McDonnell, S.M., Poulin, A., 2002. Equid play ethogram. *Appl. Anim. Behav. Sci.* 78, 263-290.
- McGreevy, P.D., McLean, A.N., Warren-Smith, A.K., Goodwin, D., 2005. Defining the terms and processes associated with equitation. In: McGreevy, P.M., McLean, A.N., Warren-Smith, A.K., Goodwin, D., Waran, N. (Eds.), 1st International Equitation Science Symposium 2005, Melbourne, Australia, 10-43.
- McGreevy, P.D., Oddie, C., Burton, F.L., McLean, A.N., 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., Sundin, M., Karlsteen, M., Berglin, L., Ternstrom, J., Hawson, L., Richardsson, H., McLean, A., 2014. Problems at the human-horse interface and prospects for smart textile solutions. *J. Vet. Behav.* 9, 34-42.
- McLean, A.N., 2008. Overshadowing: A silver lining to a dark cloud in horse training? *J. Appl. Anim. Welfare Sci.* 11, 3, 236-248.
- Meschan, E.M., Peham, C., Schobesberger, H., Licka, T.F., 2007. The influence of the width of the saddle tree on the forces and the pressure distribution under the saddle. *Vet. J.* 173, 578-584.
- Monkemoller, S., Keel, R., Hamsch, D., Muller, J., Kalpen, A., Geuder, M., Auer, J.A., von Rechenberg, B., 2005. Pliance Mobile-16HE: A study about pressure measurements under the saddle after the adjustment of the saddle fit. *Pferdeheilkunde* 21, 102-et seq.
- Morgan, M.H., 1993. *Xenophon: The Art of Horsemanship*. J.A. Allen and Company Limited, London UK.
- Morris, M., 2010. *Design of experiments: An Introduction Based on Linear Models*. Chapman & Hall/CRC Texts in Statistical Science. Boca Raton, Florida, USA.
- Nevison, C.M., Hughes, A., Cole, M.D., 2011. Variability in lower leg aids used to achieve gait transitions on a Dressage Simulator. *J. Vet. Behav.* 6, 294-294.
- Nevison, C.M., Timmis, M.A., 2013. The effect of physiotherapy intervention to the pelvic region of experienced riders on seated postural stability and the symmetry of pressure distribution to the saddle: A preliminary study. *J. Vet. Behav.* 8, 261-264.
- Nomina Anatomica Veterinaria, 2012. 5th Edition. Prepared by the International Committee on Veterinary Gross Anatomical Nomenclature ICVGAN. Authorized by the General assembly of the World Association of Veterinary Anatomists WAVA in Knoxville. http://www.wava-amav.org/nav_nev.htm [Accessed 18th July 2012].

- Normando, S., Meers, L., Samuels, W.E., Faustini, M., Ödberg, F.O., 2011. Variables affecting the prevalence of behavioral problems in horses. Can riding style and other management factors be significant? *Appl. Anim. Behav. Sci.* 133, 186-198.
- Nyikos, S., Werner, D., Muller, J.A., Buess, C., Keel, R., Kalpen, A., Vontobel, H.D., von Plocki, K.A., Auer, J.A., von Rechenberg, B., 2005. Measurements of saddle pressure in conjunction with back problems in horses. *Pferdeheilkunde* 21, 187-198.
- Ödberg, F.O., Bouissou, M-F., 1999. The development of equestrianism from the baroque period to the present day and its consequences for the welfare of horses. *Equine Vet. J. Suppl.* 28, The Role of the Horse in Europe, 26-30.
- Oka, T., Oka, K., Hori, T., 2001. Mechanisms and mediators of psychological stress-induced rise in core temperature. *Psychosom. Med.* 63, 476-486.
- Parker, M., Redhead, E.S., Goodwin, D., McBride, S.D., 2008. Impaired instrumental choice in crib-biting horses (*Equus caballus*). *Behav. Brain Res.* 191, 1, 137-140
- Peeters, M., Sulon, J., Beckers, J.F., Ledoux, D., Vandenheede, M., 2011. Comparison between blood serum and salivary cortisol concentrations in horses using an adrenocorticotrophic hormone challenge. *Equine Vet. J.* 43, 487-493.
- Peham, C., Hofmann, A., Molsner, J., Borkenhagen, B., Kuhnke, S., Baltacis, A., 2008. Forces acting on the horse's back and the stability of the rider in sitting and rising trot—a comparison. *Pferdeheilkunde* 24, 337–342.
- Peham, C., Licka, T., Schobesberger, H., Meschan, E., 2004. Influence of the rider on the variability of the equine gait. *Hum. Movement Sci.* 23, 663-671.
- Perneger, T.V., 1998. What's wrong with Bonferroni adjustments? *Brit. Med. J.* 316, 1236-1238.
- Peters, S.M., Bleijenberg, E.H., van Dierendonck, M.C., van der Harst, J.E., Spruijt, B.M., 2012. Characterization of anticipatory behavior in domesticated horses (*Equus caballus*). *Appl. Anim. Behav. Sci.* 138, 60-69.
- Physick-Sheard, P.W., Marlin, D.J., Thornhill, R., Schroter, R.C., 2000. Frequency domain analysis of heart rate variability in horses at rest and during exercise. *Equine Vet. J.* 32, 253-262.
- Preuschoft, H., Witte, H., Recknagel, S., Bar, H., Lesch, C., Wuthrich, M., 1999. The effects of various head-gears on horses. *Deut. Tierarzt. Woch.* 106, 169-175.
- Puers, R., 1993. Capacitive sensors –when and how to use them. *Sensor.Actuator. A-Phys.* 37-8, 93-105.

Pullin, J.G., Collier, M.A., Durham, C.M., Miller, R.K., 1996. Use of force sensing array technology in the development of a new equine saddle pad: Static and dynamic evaluations and technical considerations. *J. Equine Vet. Sci.* 16, 207-216.

Quick, J.S., Warren-Smith, A.K., 2009. Preliminary investigations of horses' (*Equus caballus*) responses to different bridles during foundation training. *J. Vet. Behav.* 4, 169-176.

Randle, H., Abbey, A., Button, L., 2011. The effect of different rein types on the rein tension applied when taking up a 'medium contact'. *J. Vet. Beh.: Clin. App. Res.* 6, 295-295.

Rothman, K.J., 1990. No adjustments are needed for multiple comparisons. *Epidemiology* 1, 43-46.

Schaefer, A.L., Stewart, M., Webster, J.R., Cook, N.J., Colyn, J.J., Lepage, P., Church, J.S., Haley, D.B., 2006. Objective measurement of pain and fear in cattle using infrared thermography. Proceedings of the International Society of Applied Ethology, North American Regional Meeting. Vancouver, Canada: The University of British Columbia; p. 55. [http://www.Appliedethology.org/isaemeetings_files/2006%20Canada-USA%20Regions%20ISAE%20Proceedings.pdf].

Schmidt, A., Möstl, E., Aurich, J., Neuhauser, S., Aurich, C., 2009. Comparison of cortisol and cortisone levels in blood plasma and saliva and cortisol metabolite concentrations in faeces for stress analysis in horses. In: 5th International Equitation Science Conference. (Eds): P. McGreevy, A. Warren-Smith, C. Oddie. Sydney, Australia 53.

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

Schmidt, A., Möstl, E., Wehnert, C., Aurich, J., Müller, J., Aurich, C., 2010b. Cortisol release and heart rate variability in horses during road transport. *Horm. Behav.* 57, 209-215.

Staller, L., 2005. Understanding analog to digital converter specifications. *Embedded Systems Design*. embedded.com. [accessed 20th Sept 2013.]
<http://www.embedded.com/design/embedded/4025078/Understanding-analog-to-digital-converter-specifications>

Starke, S.D., Willems, E., May, S.A., Pfau, T., 2012. Vertical head and trunk movement adaptations of sound horses trotting in a circle on a hard surface. *Vet. J.* 193, 1, 73-80.

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

- Stewart, M., Stafford, K.J., Dowling, S.K., Schaefer, A.L., Webster, J.R., 2008. Eye temperature and heart rate variability of calves disbudded with or without local anaesthetic. *Physiol. Behav.* 93, 789-797.
- Strand, S.C., Tiefenbacher, S., Haskell, M., Hosmer, T., McDonnell, S.M., Freeman, D.A., 2002. Behavior and physiologic responses of mares to short-term isolation. *App. Anim. Behav. Sci.* 78, 145-157.
- Suwala, M., Jezierski, T., Gorecka-Bruzda, A., 2012. Localisation of artefacts in two models of equine heart rate monitors –a pilot study. In: Randle, H., Waran, N. Williams, J., (Eds.), 8th International Equitation Science Symposium 2012, Edinburgh, United Kingdom, 90.
- Task Force of the European Society of Cardiology, t.N.A.S.o.P.E., 1996. Heart rate variability: standards of measurement, physiological interpretation, and clinical use. *Circulation* 93, 1043-1065.
- Tateo, A., Maggiolino, A., Padalino, B., Centoducati, P., 2012. Behavior of artificially suckled foals. *J. Vet. Behav.* 7, 1, 33-42.
- Taylor, K.D., Mills, D.S., 2006. The development and assessment of temperament tests for adult companion dogs. *J. Vet. Behav.* 1, 3, 94-108.
- Thivierge, M., Leger, L., 1989. Critical review of heart rate monitors. *CAHPER J.* 55, 26-31.
- Valera, M., Bartolomé, E., Sánchez, M.J., Molina, A., Cook, N., Schaefer, A., 2012. Changes in eye temperature and stress assessment in horses during show jumping competitions. *J. Equine Vet. Sci.* 32, 827-830.
- Visser, E.K., van Reenen, C.G., Hopster, H., Schilder, M.B.H., Knaap, J.H., Barneveld, A., Blokhuis, H.J., 2001. Quantifying aspects of young horses' temperament: consistency of behavioral variables. *Appl. Anim. Behav. Sci.* 74, 241-258.
- Visser, E.K., Reenen, C.G. van, Werf, J.T.N. van der, Schilder, M.B.H., Knaap, J.H.; Barneveld, A., Blokhuis, H.J., 2002. Heart rate and heart rate variability during a novel object test and a handling test in young horses. *Physiol. Behav.* 76, 289 – 296.
- von Borstel, U.U., Keeling, L.J., Duncan, I.J.H., 2005. Transfer of nervousness from the rider to the horse, 39th International Congress of the International Society of Applied Ethology, Sagamihara, Japan, 84.
- von Borstel, U.U., Duncan, I.J.H., Shoveller, A.K., Merkies, K., Keeling, L.J., Millman, S.T., 2009. Impact of riding in a coercively obtained Rollkur posture on welfare and fear of performance horses. *Appl. Anim. Behav. Sci.* 116, 228-236.

von Peinen, K., Wiestner, T., Bogisch, S., Roepstorff, L., van Weeren, P.R., Weishaupt, M.A., 2009. Relationship between the forces acting on the horse's back and the movements of rider and horse while walking on a treadmill. *Equine Vet. J.* 41, 285-291.

von Peinen, K., Wiestner, T., von Rechenberg, B., Weishaupt, M.A., 2010. Relationship between saddle pressure measurements and clinical signs of saddle soreness at the withers. *Equine Vet. J.* 42, 650-653.

Warren-Smith, A.K., Curtis, R.A., Greetham, L., McGreevy, P.D., 2007. Rein contact between horse and handler during specific equitation movements. *Appl. Anim. Behav. Sci.* 108, 157-169.

Werner, D., Nyikos, S., Kalpen, A., Geuder, M., Haas, C., Vontobel, H.-D., Auer, J.A., von Rechenberg, B., 2002. Druckmessungen unter dem Sattel: Eine Studie mit einem elektronischen Sattel-Messsystem (Novel GmbH). *Pferdeheilkunde* 18, 125–140.

Winkelmayr, B., Peham, C., Frühwirth, B., Licka, T., Scheidl, M., 2006. Evaluation of the forces acting on the back of the horse with an English saddle and a side saddle at walk, trot and canter. *Equine Vet. J. Suppl.* 36, 406-410.

Yarnell, K., 2012. An investigation into the use of infrared thermography as a tool to assess the physiological stress response in the horse. PhD thesis. Nottingham Trent University.

Yarnell, K., Hall, C., Billett, E., 2013. An assessment of the aversive nature of animal management procedures using behavioral and physiological measures. *Physiol. Behav.* 118, 32-39.