The validation of the Composite Pain scale (CPS) for acute colic in horses

and

the influence of horse personality in responses to pain scored using CPS

<u>Author</u>:

M.R. van Gent

3382265

August – November 2013

<u>Supervisors:</u>

Dr. J.P.A.M van Loon

Dr. M.C. van Dierendonck

Table of contents

Abstract

Currently no standard pain severity scale for acute abdominal pain (colic) in horses is in clinical use. In this study 25 colic patients and 25 control horses have been scored by two independent observers using the CPS after admittance in referral clinic. Sensitivity and specificity of CPS for differentiation between colic horses (pain) and control horses (no pain) are resp. 96% and 84%. Statistics on CPS show a high inter-observer reliability significant on a $P<0,001$ level. The CPS scores were compared to the scores of the same horses by means of a Visual Analogue Scale (VAS). The VAS scores are determined by two independent experienced veterinarians. VAS has been used as current standard in this experiment. Sensitivity and specificity for this scale are lower than those obtained for the CPS resp. 57% and 60%. Furthermore the inter-observer reliability showed less correlation than for CPS. A low significant correlation is demonstrated between CPS and VAS ($P=0,021$). CPS seems to be more useful than VAS to assess pain in horses having acute colic. The Composite Pain Scale (CPS) is a promising tool for assessing abdominal pain in horses.
Besides the validation of the CPS, the influence of personality of individual horses on the CPS scores is investigated in this study. Horse personality is assessed by questionnaires completed by handlers of the horse and by a few behavioral tests performed by two students on the second day after the horse arriving at the clinic. It was expected that proactive horses would score higher on CPS based on more behavioral expression of pain compared to reactive horses. Not enough data were collected to statistically analyze these data. More research should be done after this subject.

Introduction

Pain

In the last decade there has been growing interest in the recognition and management of pain responses in veterinary medicine. Pain can be defined as a sensory and emotional experience generated by the activation of nociceptors associated with actual or potential tissue damage. This experience causes a physiological and behavioral response to reduce or avoid damage, to reduce the likelihood of recurrence and to promote recovery (Molony and Kent, 1997; Muir, 2010; Taylor et al., 2002).

Noxious (pain-producing) stimuli activate nociceptors located throughout the body. These stimuli trigger a system of transduction (detection), transmission, modulation, projection and central processing (Figure 1). These neural processes together are called nociception. The nociceptors convert the noxious stimuli to electrical impulses, which in case of gastro-intestinal pain are transmitted to the spinal cord by unmyelinated first order sensory nerve fibers called C fibers. These C fibers cause the slower onset of dull pain. Because the C fibers travel with nerves of the autonomic system autonomic physiologic responses like hyperpnoea, tachycardia and sweating are associated with visceral pain. After modulation in the dorsal horn of the spinal cord, the second order neurons project the impulses on the brain. In the brain physiologic and emotional responses are generated, which lead to the expression of pain by the individual(Muir, 2010).

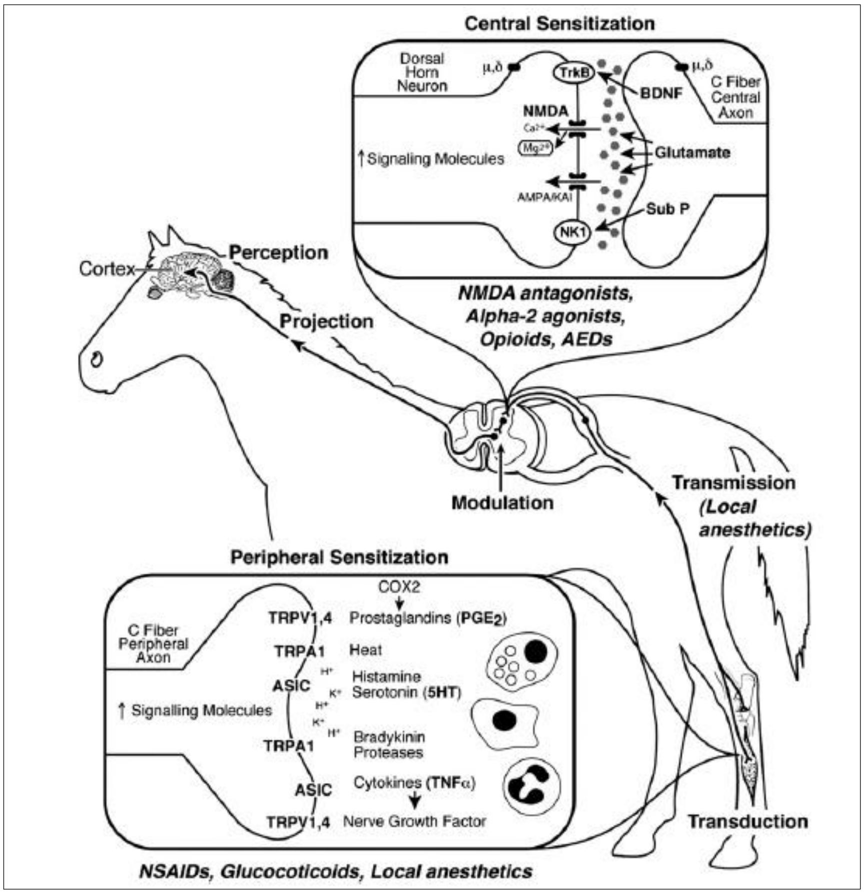

Figure 1; nociceptive pathways (Muir, 2010)

Because this response is multidimensional it can be challenging to objectively assess pain in non-verbal individuals like horses. In horses it can be even harder to assess pain than in non-verbal individuals, because horses are flight animals and showing pain is a sign of weakness. This results in sometimes little to no pain expression even when the horse is in heavy pain. Even for experienced veterinarians it is difficult to distinguish if a horse is in light or severe pain. Pain assessment in horses depends on interpretation of, sometimes subtle, changes in behavior and physiology (Price et al., 2003; Taylor et al., 2002; van Loon et al., 2010).
A scale that makes it possible for clinicians (or even owners) to assess pain more evidence-based and objectively, would contribute to the animal welfare.

Pain scales

The first scales used to assess pain in animals were based on human pain scales. Because people are able to describe what they feel, the most objective way is to rate the pain based on the patient's own opinion. The patient is asked to verbally rate his or her pain on a scale from 0 to 10 where 0 is 'no pain' and 10 is 'worst pain possible' using the Numeric Rating Scale (NRS) This scale is a categorical scale which means that every number from 0 to 10 stands for a specific description of the pain (no pain, little pain, moderate pain, heavy pain etc) (Bijur et al., 2003). Another way to use the patient's opinion is to ask to place a mark on a line of 10 cm, with at the ends 'no pain' and 'worst pain possible', which indicates their severity of pain by means of a Visual Analogue Scale (VAS). The VAS is a continuous scale, so there are no categories and a mark can be placed everywhere on the line that fits to the intensity of the pain (Bijur et al., 2001). Facial Pain Scale (FPS) is commonly used to assess children's pain, because they are limited in verbal communication and associative thinking. Children are asked to point at a drawing of a face, which indicates the pain they feel (Figure 2) (Hicks et al., 2001).

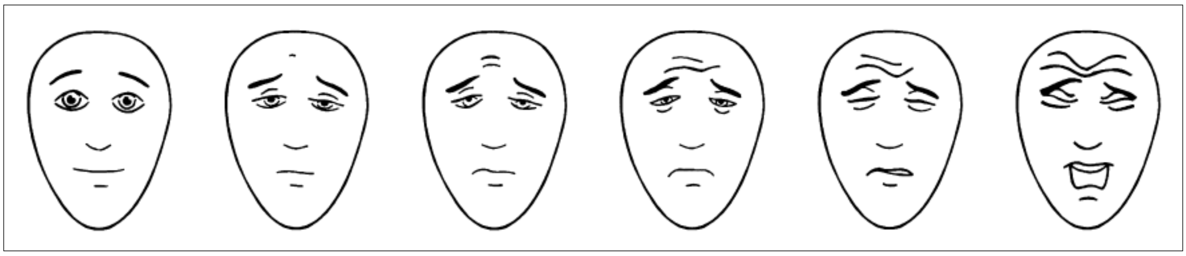

Figure 2 Human Facial Pain Scale; drawings express the degree in pain from 'no pain' to 'worst pain possible'(Hicks et al., 2001)

In non-verbal humans (like babies or disabled) and animals it is not possible to use the verbal pain indication of the patient itself. Therefore pain can only be observed with indicators like behavioral and physiological changes (Molony and Kent, 1997).
Besides the use of the VAS for rating the pain experience of humans, the VAS can also be used by an observer to rate the pain of an animal. The observer places a mark on the line corresponding to their interpretation of the patient's pain intensity. Because the score is dependent of the capability of the observer to determine pain, this is a subjective way of

assessing pain and the reliability of this scale is low. Behavior-based scales have shown to be more reliable (Hielm-Björkman et al., 2011; Sutton et al., 2013).
Recent studies have developed more objective multi-dimensional pain scales, which can relatively accurately measure pain in horses based on behavior and physiology. Because every animal has its own way of expressing pain (for example prey animals vs. predators) most scales are species specific (Taylor et al., 2002). Although there are some non-specific indicators of pain, most scales are designed for a specific type of pain (Ashley et al., 2005). Pritchett et al.(2003) revealed specific abdominal pain indicators based on behavior and physiologic after surgery and then Graubner et al. (2011) designed a Post Abdominal Surgery Pain Assessment Scale (PASPAS) to measure pain using those indicators. Other scales have been developed to assess orthopedic pain in horses (Bussieres et al., 2008; Price et al., 2003).

Because animals are capable of showing an innate and adaptive facial expression of emotion like pain, there are pain scales based on behavior and facial acting codes. One of the first animal facial expression scales was the Mouse Grimace Scale designed for laboratory animals (Langford et al., 2010). Very recently some facial expression scales have been designed for horses. Love et al. (2011) showed that horses display a 'pain face' using facial acting codes. Pain changes the distance between measuring points at the head what causes a grimace (Love et al., Conference abstract. 2011). A Facial Expression Pain Scale (FEPS) has been designed for evaluating post-castration pain in horses (Dalla Costa et al., 2014).

One of the latest and possibly most reliable pain evaluation tools is the Composite Pain Scale (CPS). The CPS is a composite numerical rating scale incorporating multiple factors to make the outcome more reliable. It was originally designed by Bussieres et al. (2008) to assess orthopedic pain in horses, but, slightly adapted, proved to be useful for rating non-orthopedic surgical and non-surgical acute visceral or somatic pain as well (Bussieres et al., 2008; van Loon et al., 2010; Van Loon et al., 2014). To be a reliable and valid scale, the scale should give a repeatable interpretation from one evaluator to another, it should be sensitive to score pain and closed to misinterpretation. This means inter-observer correlation should be high, there should be a high sensitivity and specificity and there should be an agreement between scoring systems used (Ashley et al., 2005; Bussieres et al., 2008).

Coping styles

Between animals there are different ways of reacting on a stressor like pain, every animal has its own coping style. Coping style is a term that is used to describe a consistent behavioral and physiological effort to master a certain situation within a certain individual (Koolhaas et al., 1999). Coping styles can be divided in subcategories: with at the extremes proactive and reactive. With proactive behavior the animal actively or aggressively tries to control the stress, while reactive animals have a more flexible and adaptive way of reacting on environmental challenges. These ways of responding are not rigid and will not be the same in different situations in one individual. But evidence has been found in different species that the behavioral and physiological response of individual animals to a specific stressor is consistent over time (Ijichi et al., 2013; Koolhaas et al., 1999; Mills, 1998). Because the CPS is based on the behavior and physiology of the animal, it is important to know the animal's coping style so it can be assessed that the scale does reflect pain and not only the coping style of the animal (Ashley et al., 2005; Ijichi et al., 2013; Price et al., 2003; van Loon et al., 2010). Which coping style the animal shows is dependent of the personality of that individual. Personality is described as 'the biologically based behavioral tendencies of an individual' (Mills, 1998). Personality is partly influenced by nurture; individuals could show different behavior in reaction to a certain stimulus because they may have had an experience that altered their behavior. On the other hand personality is partly heritable; species bred for their hardiness (Shetland ponies, Friesian horses) would show less pain behavior than species bred for their speed (Thoroughbreds) (Taylor et al., 2002).

Personality assessment

There are different ways to assess personality of, in this study, the individual horse. Personality can be determined by the reaction of an individual on a series of validated practical tests, like a novel object test, open field test, passive and active human approach test etc. To perform each of these separately is time consuming, requires the skills of a trained scientist and often measures just one personality trait (Ijichi et al., 2013). Another way of determining the personality is subjective assessment: (Ijichi et al., 2013) showed that personality traits of an animal are scored by a known care-giver who frequently interacts with the animal gives a reliable result. Most published and used horse personality questionnaires like those of Lloyd et al. (2007) & (2008), Ijichi et al. (2013) and Momozawa et al. (2003)are based on scoring of behaviorally defined adjectives (BDA's). These BDA's are rated by a handler of the horse on a Likert-type scale: 1-3, 1-5 or 1-7. Principle component analysis (PCA) or Factor analysis (FA) identifies a number of personality components/factors based on grouping of the BDA's. These personality components are labeled with a term that fits the combination of BDA's (e.g. sociability, excitability etc.)(Ijichi et al., 2013; Lloyd et al., 2007; Lloyd et al., 2008; Momozawa et al., 2003). There is strong evidence that the use of a questionnaire of this type is a reliable method to assess horse personality (Lloyd et al., 2007). Combining practical tests and questionnaire will probably give the most reliable outcome (Ijichi et al., 2013).

Research goal & hypothesis

Research goal – CPS validation

The goal of this study was to construct a new CPS for horses with pain due to acute colic. This scale was internally validated by comparing CPS to the current standard VAS using inter-observer reliability measurements in both scales and sensitivity/specificity for differentiating between colic horses (pain) from control horses (no pain).

Hypothesis

CPS is more reliable to measure pain in horses with acute abdominal pain compared to VAS.

Because CPS is a composite scale of precisely described behavioral and physiological elements, the sum of these scores would be more accurate than just one score (VAS) based on experience of one person (subjective). This will likely result in a higher inter-observer reliability and a higher sensitivity and specificity for pain due to acute colic.

Research goal – influence of personality

The second goal of this study was to determine if the personality of the horse has an influence on the pain score of the patient with acute colic, using CPS. And if so, to determine which factors had a correlation. The results can be used to assess if the CPS needs to be clarified for the different factors or needs to be adjusted.

Hypothesis

There could be a difference in pain score using CPS between horses that have a different personality while suffering from the same type of acute colic.
Assuming that horses with a higher level of anxiousness and excitability have a more active coping style than horses that have a low level on these two personality components, they were expected to show more (frequent of severe) signs of colic behavior like kicking at the abdomen, pawing at the floor and reacting on palpation of the painful area. Because behavior is a large part of the CPS, this difference in behavior based on personality would result in a higher score on CPS while having the same stressor (pain due to acute colic).

Materials & Methods

Animals

In this study horses with acute visceral pain (colic) admitted to the Department of Equine Science of the Faculty of Veterinary Medicine in Utrecht between 19/08/2013 and 6/10/2013 were included. This patient group was composed of 17 horses: 7 geldings and 10 mares aged 3-22 years. Foals less than one year old, mares caring a foal and stallions were excluded on account of different behavior and/or other physiological parameters compared to adult mares and geldings. 8 horses (7 mares and 1 gelding) aged 4-18 years old housed at the clinic for educational purposes were used as control group. These horses were selected on not having any (abdominal) pain.

Besides these test and control horses, data collected by previous students in the CPS study were also used, including a patient group of 8 horses (5 geldings, 3 mares) aged 4-20 years old and a control group of 18 horses; 10 mares owned by the Reproduction department of the clinic aged 2-13 years and a group of 7 geldings and 1 mare aged 5-16 years that came to the farrier of the clinic for change of shoeing. These horses were chosen on the same selection criteria as we did in the recent study. One control mare was excluded from the database because retrospectively seen this horse had serious tooth problems and a huge worm infection when scored, so she was certainly not free of pain. An overview of patient and control groups is displayed in table 1 and 2.

Patient-group	Number (N)	gender		Age (years)
Recent study	17	7♂	12♀	3-22
Database	8	5♂	3♀	4-20
Total	**25**	**12♂**	**15♀**	**3-22**

Table 1

Control-group		Number (N)	gender		Age (years)
Recent study	education	8	1♂	7♀	4-18
Database	reproduction	9	0♂	9♀	2-13
	farrier	8	7♂	1♀	5-16
Total		**25**	**8♂**	**17♀**	**2-18**

Table 2

Horses used in the study divided in patient (N=25) in table 1 and control group (N=25) in table 2. The horses from the database are scored in a former study. The distribution in gender and age is showed for every group.

Pain scales

CPS +FEPS

A modified version of the CPS of Bussieres et al. (2008), van Loon et al. (2010) & Van Loon et al. (2014) adjusted by the previous students in the CPS study, was used to assess a pain score of the colic horse (Annex I). This pain scale is a composite numerical rating scale consisting of 14 elements grouped in 3 factors; behavioral, physiologic or interactive. Each element was rated a score 0 to 3, where 0 representing a normal not painful condition and 3 represents the most abnormal painful value. The total score could range from 0 to 42. The Facial Expression Pain Scale was designed by the previous students as well, based on published pain characteristics of the equine head (Annex II)(Ashley et al., 2005; Love et al., Conference abstract. 2011; Dalla Costa et al., 2014; Taylor et al., 2002). The possible scores per element were 0, 1, 2 or 3 where 0 represented a normal not painful condition and 3 represented the most abnormal painful value. Some elements were just scored 0 or 2 when respectively absent or present. The total score could range from 0 to 18.

<u>The procedure:</u>

After entry in the clinic the treating veterinarians did their work. At a suitable moment the owners were asked for allowing us to observe their horse and use the results in this study. The owners were informed about the application of the pain scoring system for their horse. After consent from the owners and at a suitable moment according to the treating veterinarians, the horse was placed in a colic box and left alone for (at least) 5 minutes to score the pain using the CPS. At the same time the behavior was recorded using a fixed camera in the colic box. After 5 minutes the muzzle (which prevented the horse to eat) was removed; one of the observers hold the head more or less fixed and the head was filmed for 1 minute frontal and slightly sideways / lateral using a mobile phone. About 20 seconds after the start of the filming a few times a noise (clicking with the tongue or fingers) was made to attract attention of the horse and to record its reaction. This 1-minute film clip was scored afterwards by means of the FEPS by the same observers. A pair of observers performed the scoring of the CPS and FEPS simultaneously but independently. Obviously these observers could not be blinded for the reason of admittance or treatment of the case. The practical arrangement for the 5-minute observation is shown in figure 3.

To monitor the pain and notice the changes in pain in time the horses were measured repeatedly at *T = (0)*, *T = (1st morning)* and *T = (2nd morning)* after arrival. If the patient received analgesic drugs before arrival or the pain management before arrival was unknown, the horse was also scored at *T = (2-3)* hours after arrival. The patient was also scored at *T = (1)* hour after it was placed in the box in case it was necessary for animal welfare to treat it with NSAIDS/opiates before placing the horse in the colic box (Annex I). During scoring all horses resided in the same (colic) box, to minimize any environmental influences.

The procedures were exactly the same for the control horses. The control horses were scored only once at *T = (0).*

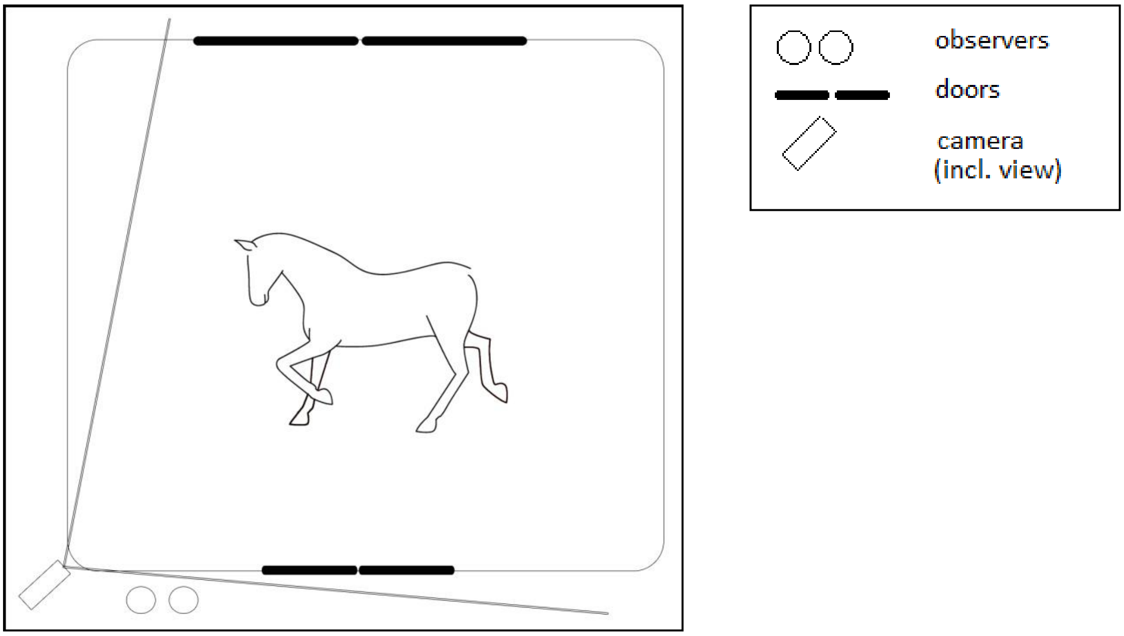

Figure 3
Practical arrangement at the colic box for the 5-minute scoring and filming

VAS

All the horses were recorded simultaneously while scoring the CPS in the colic box. The VAS was used as the standard to compare with the CPS. Therefore two independent veterinarians also scored the recorded video's. In order to do so all clips were shortened from 5 minutes to only the first minute that was scored. This adaptation was performed for practical reasons, because it would take too much time for the independent veterinarians to score a 5 minute film. All the 1-minute film clips (patient and control group) were randomized, blinded and saved on an USB-stick. These 1-minute film clips were shown individually to two equine veterinarians. These two veterinarians were both doing an internship at the Equine University Clinic of Utrecht. They have scored pain severity of the colic horses according to a VAS by indicating an X on an undivided line of 10 cm (Annex III). The scoring options ranged continuously from 0 to 10, where 0 meant 'no pain at all' and 10 represented 'worst pain possible'(Hielm-Björkman et al., 2011; Sutton et al., 2013). Afterwards the scores were measured in millimeters (0-100).

Use of data

The data-set used consisted 72 mean CPS scores of all test moments of every horse (test and control group, database plus recent study) scored by two observers. Because of technical problems it was not possible to produce film clips of the first 7 patients of this study, so 57 1-minute film clips remained to be scored by means of a VAS.
In the data analysis the VAS scores of just one observer (observer 1) were used in the statistical calculation because of the high inter-observer variability within the VAS scorers. Previous students who created the database used the data of the same observer what makes it comparable with the data of this study.
All 72 test moments of CPS and FEPS and 57 1-minute film clips of VAS were scored by two independent observers. The inter-observer correlation coefficient of every scale used is

determined by Spearman's rho test. All results were divided in results of data from former study (observers Marianne/Larissa), results of data of this study (observers Anne-Marie/Marlijn) and results of the combined data.
Mann Whitney U test was used to compare the scores on different scales of the patient group (32 measurements at different times) to the control group (25 measurements at *T= (0))* The differences between colic and control group are shown in a box plot for every scale.
For determining sensitivity, specificity, positive predictive value and negative predictive value the 'acute phase' scores were used. 'Acute phase' scores meant the scores of the first time the patient was rated. This comprises the scores of *T = (0),* or *T= (1)/T = (2-3)* if it was not possible to get a rating on *T= (0)*. The 'acute phase' patient group consisted of 24 horses for CPS and FEPS and of 14 horses for VAS. The control group consisted of 25 horses for all of the comparisons (recent study plus database).
To distinguish no pain/no colic' from 'pain/colic' two cut off scores were determined: within CPS all scores ≤5 represented 'no pain' (negative) and >5 represented 'pain' (positive). The cut off scores for CPS were determined as 5 < X ≤ 11 for 'moderate pain' and > 11 for 'heavy pain'. For FEPS the cut off score was <3 negative and ≥3 positive; within the VAS a score of ≤1 represented negative and >1 a positive. The scores of CPS and FEPS were combined and treated the same way as the separate scores. The cut off scores for the combination were ≤8 negative and >8 positive.
The sensitivity, specificity, positive predictive value and negative predictive value were calculated by the use of the following table:

scoring system	T0/T1/T2-3			
	patient/colic	control		
cut- off score 'no pain'	a	b	negative predictive value	(b/a+b) x 100
cut- off score 'pain'	c	d	positive predictive value	(c/c+d) x 100
	sensitivity	specificity		
	(c/a+c) x 100	(b/b+d) x 100		

Sensitivity shows the percentage of colic horses that is also scored above the cut-off score for 'pain'(positive), so it is a parameter for false negative scores. Specificity shows the percentage of control horses that is also scored below the cut-off score for 'no pain'(negative), so it is a parameter for false positive scores. The negative predictive value shows the percentage of horses that are scored below the cut-off score for 'no pain' that are really negative (controls). The positive predictive value shows the percentage of horses that are scored above the cut-off score for 'pain' that are really positive(patient /colic horses).
Z-value was used to distinguish 'specific' colic pain elements from the other pain elements scored by CPS. For each separate element the total frequency scored for every patient at *T= 0* was determined and the z-value was calculated by the formula: (patient frequency – control frequency)/(patient frequency + control frequency). All elements that resulted in a score above +0,5 were labeled as specific for (colic) pain.

Questionnaire & behavioral tests

Questionnaire

At least one handler (preferably the handler who knew the horse the best) of each patient was asked to complete a validated personality questionnaire by Emma Creighton in (Ijichi et al., 2013). The questionnaire was translated in Dutch (Annex IV). To obtain another indication of the personality of the horse the treating student (group) at the clinic was asked to complete the questionnaire as well. Instructions on how to complete the questionnaire were provided. The raters were asked not to discuss the answers with each other to obtain the most objective outcome.

The questionnaire was composed of 15 continuous lines of 10 cm with at each end opposite behaviorally defined adjectives (BDA's). These BDA's are placed in randomized sequence at the endpoints of the lines. All BDA's are based on five factors: agreeableness (A), neuroticism (N), extroversion (E), gregariousness to people (F4), gregariousness to horses (F5)(Ijichi et al., 2013). Afterwards the marks on the lines were subsequently divided in 5 (equal) parts to give the mark a numerical rating from 1 to 5, where 5 stands for positively correlated to the factor and 1 means negatively correlated to the factor. The score code of the original version was used to determine to which factor the BDA belongs. 'R' stands for 'reverse', so the rating for that line was 5-1. The last 3 questions were already numerically rated.

Behavioral tests

At *T= (2nd morning)* each conservatively treated patient was tested on four personality traits: attentiveness, reactivity, sensitivity and fearfulness. Each personality trait was tested by a non-invasive test in the colic box. The test-protocol was based on several types of tests that have been used to quantify animal personalities and have shown to be consistent over time by other authors: (Christensen et al., 2008; Ijichi et al., 2013; Unknown, 2012) All tests were conducted identically and in the same order. Fearfulness was tested by a 'novel arena test' (first minute in the box) and by the reaction on a 'negative sound' (loud whistle from a mobile phone). Reactivity was tested by a 'human approach test' and the reaction the 'garbage bag test' as shown in figure 4 (removing a garbage bag from underneath a food bucket). Attentiveness was scored by a 'novel object test' (showing a colored wallet) as shown in figure 5; the reaction on a 'positive sound' (food pellets in a bucket) and the recovery of the 'garbage bag test' (time until approaching and eating from the bucket again). Sensitivity was scored by the 'skin sensitivity test'. The traits were scored on a score 1-3 or 1-5. The test protocol is described in Annex V.

Figure 4
Garbage bag test; horse eating from a bucket.

Figure 5
Novel object test; reaction of the horse on a colored wallet.

Statistical analysis

Pain scales

The reliability of the scoring systems was determined by testing the inter-observer correlation coefficient. Spearman's rho analysis was used to determine inter-observer correlation of the CPS, FEPS and VAS between the two independent raters. To determine the validity of the scales the correlation between different scoring systems was tested by means of Spearman's rho test. Sensitivity and specificity of the different scales was used to assess the validity of the scales. The value of each element of the CPS separately is determined by calculating the face value (z – value). Mann Whitney U test was used to test if there was a significant difference between scores of colic patient vs. control horses for every scale.
All data were processed in Microsoft Excel 2000 and statistical analysis was performed using IBM SPSS statistics 20. Significance for all statistical analysis was set at $P<0,05$

Personality assessment

All factors were grouped in components using Principal Component Analysis (PCA) with Varimax rotation. Components were extracted and were usable if the Scree plot was >1. Factors loaded strongly to the component when having a value >0.6 or <-0.6. Components were given a label based on the most strongly loaded factor (Ijichi et al., 2013). For every horse the loading to each component was determined by SPSS. Then every horse was grouped into one of the components that had the highest positive correlation or the least low negative correlation. The scores on CPS of the different labeled groups of horses were compared using Mann Whitney U test and shown in a box plot.

Results

Inter-observer reliability

The correlation of the two independent scores of all scales is shown in table 3.

Inter-observer correlation

	N	Correlation coefficient (r)	Significance (P)
CPS	72	0,942	<0,001
FEPS	72	0,844	<0,001
VAS	57	0,336	0,011

Table 3
Inter-observer correlation of the different scales determined by Spearman's rho test

In the figures 6, 7 and 8 the scores of the two observers of the first study (blue) and the two observers of the recent study (red) are differentiated. The green line represents the mean correlation of all observers and the black line shows when the correlation is perfect: coefficient of 1 (line y = x). When comparing the graphs, CPS and FEPS show a better correlation than VAS. The inter-observer correlation of CPS is the highest of the three scales.

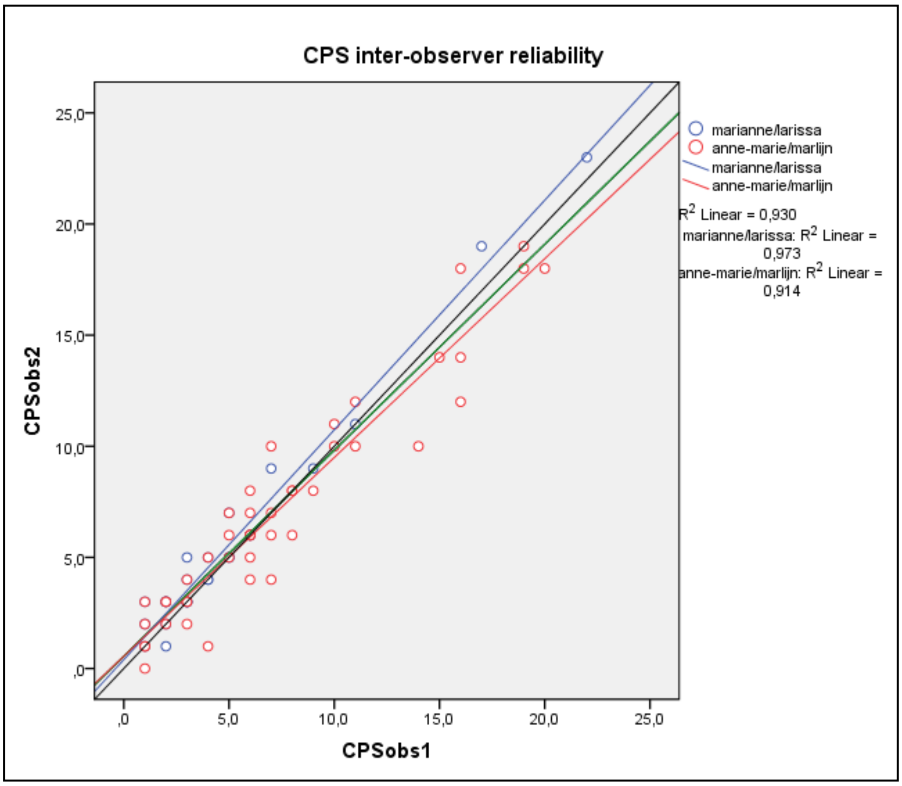

Figure 6
Inter-observer reliability of CPS; CPSobs1 = blue: obs1 of study 1, red: obs1 of study 2, green: combined. CPSobs2= blue: obs2 of study 1, red: obs2 of study 2, green: combined. Blue line: correlation of former study, Red line: correlation of this study, Green line: correlation of combined data (former and recent study). Black line: y=x, optimal correlation (1).

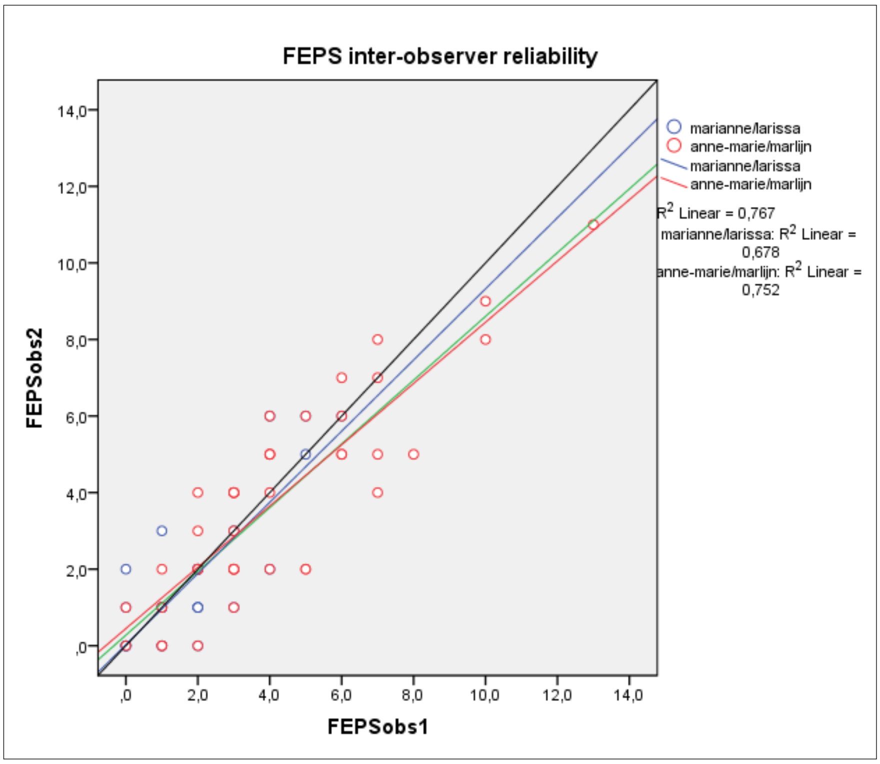

Figure 7
Inter-observer reliability of FEPS; FEPSobs1 = blue: obs1 of study 1, red: obs1 of study 2, green: combined. FEPSobs2= blue: obs2 of study 1, red: obs2 of study 2, green: combined. Blue line: correlation of former study, Red line: correlation of this study, Green line: correlation of combined data (former and recent study). Black line: y=x, optimal correlation (1).

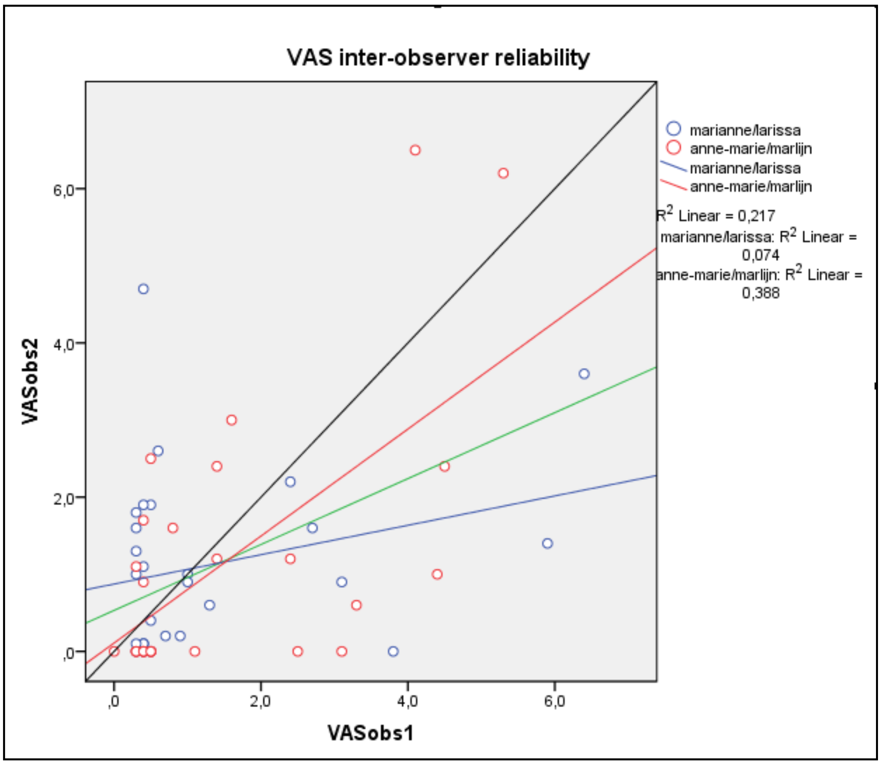

Figure 8
Inter-observer reliability of VAS; VASobs1 = blue: vet 1, red: vet 1, green: combined. VASobs2= blue: vet 2a, red: vet 2b, green: combined. Blue line: correlation of study 1, Red line: correlation of this study 2, Green line: correlation of combined data (study 1 and 2). Black line: y=x, theoretical optimal correlation (1).

Learning curve

Figure 9 shows the differences in scores on *T = (0)* between the two observers of this study. The divergence in the total score on CPS of the two observers for each horse scored on *T= (0)* was consistent over time.

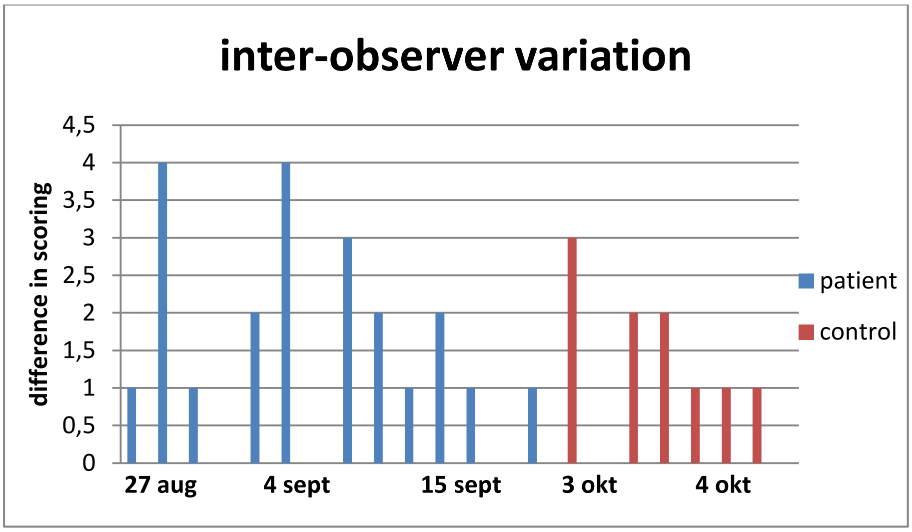

Figure 9
Inter-observer variation: difference between two total scores on CPS at T= (0) is displayed for all patient horses (N= 14) (blue bars) and for all control horses (N=8) (red bars). No bar means equal scores (0).

Correlation different scoring systems

Data of the previous students (blue) are differentiated from the data of the students of recent study (red). The green line shows the correlation between CPS and VAS of the total data-set (57 measurements) of former and recent study (Figure 10). The correlation coefficient and significance are shown in table 4. There was a significant correlation between CPS and VAS scores, but the correlation coefficient was low.

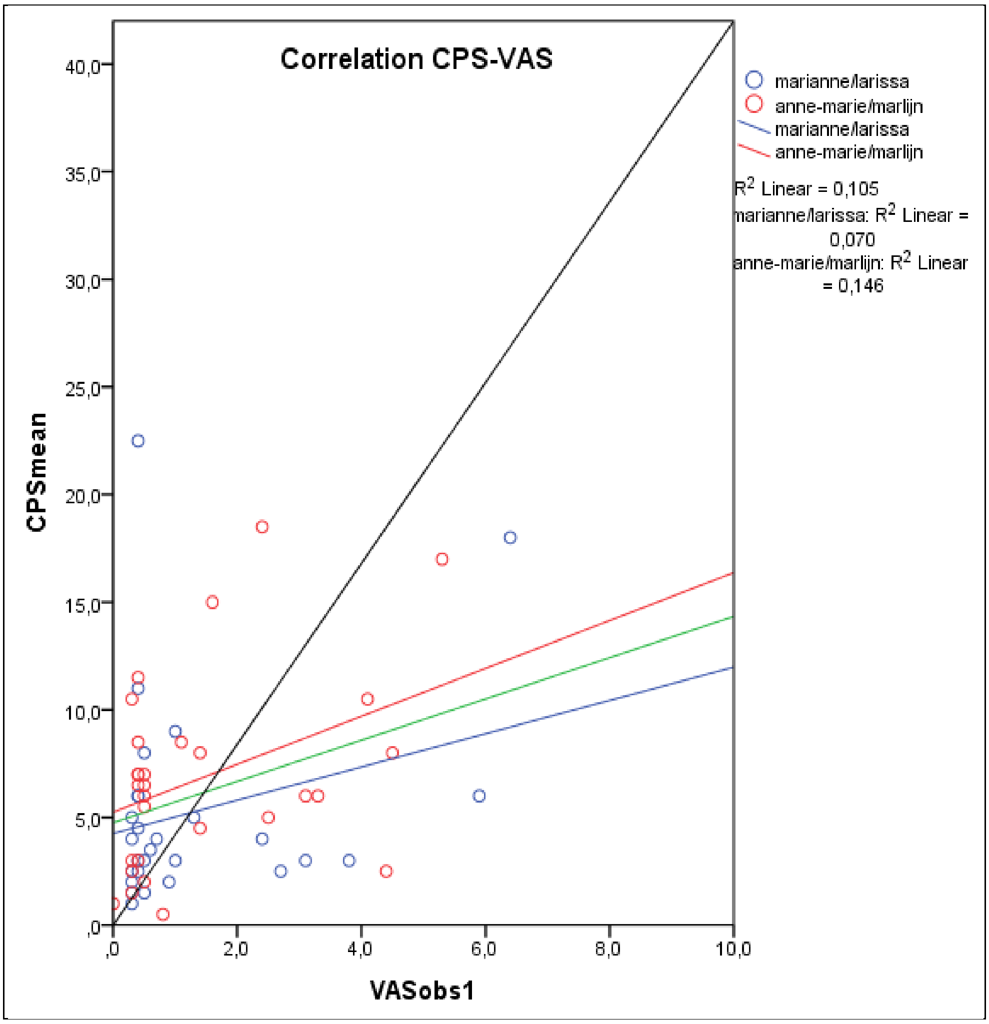

Figure 10
Correlation between CPS and VAS; CPSmean = blue: mean of scores study 1; red: mean of scores study 2, green: mean of both studies. VASobs1 = blue: scores of study 1 by vet 1, red; scores of study 2 by vet 1, green: mean of all scores of vet 1. Blue line: correlation of study 1, Red line: correlation of study 2: Green line: correlation of combined data (study 1 and 2). Black line: indication of a theoretical optimal correlation between CPS and VAS.

Correlation coefficient

	N	Correlation coefficient (r)	Significance (P)
CPS-VAS	57	0,306	0,021

Table 4

Sensitivity/specificity

Colic vs. control scored by the CPS or the FEPS showed a significant outcome (P<0,05), while colic vs. control scored by means of a VAS was not significant (P>0,05) as displayed in figure 11. Both CPS and FEPS were even significant on a 0,001 level.

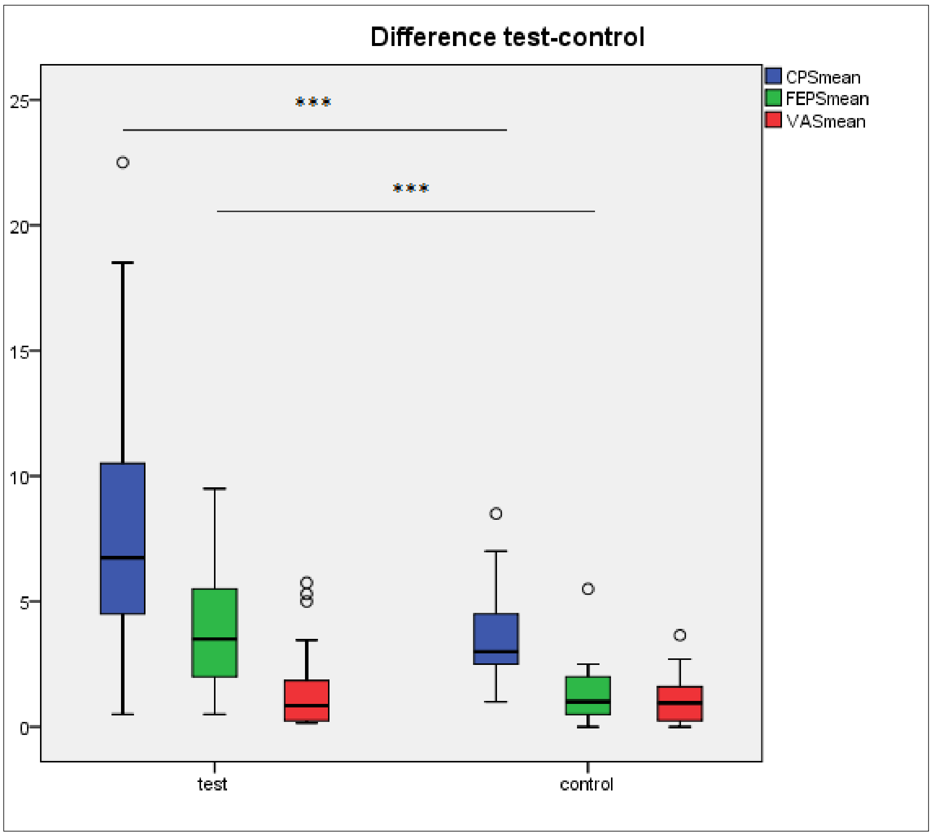

Figure 11
Median scores of both observers for colic patients (test); control horses (control). Blue boxes were CPS scores (P<0,000), Green boxes were FEPS scores (P<0,000), red boxes were VAS scores (P=0,717).

CPS and FEPS scored high on sensitivity and specificity for the differentiation between colic horses (pain) and control horses (no pain), but the combination of the two tests resulted in the highest sensitivity and specificity and highest positive and negative predictive value (table 5). The VAS test scored quite low on sensitivity (57%) and specificity (60%) as shown in table 5.

Sensitivity & specificity of different scoring systems for controls vs. patients

	Sensitivity	Specificity	Pos. pred. val.	Neg. pred. val.
CPS	96%	84%	85%	95%
FEPS	87,5%	96%	95,5%	89%
CPS + FEPS	96%	92%	92%	96%
VAS	57%	60%	56%	71%

Table 5
Differences in sensitivity, specificity, positive predictive value and negative predictive value. Sensitivity: the percentage of colic horses that is also scored above the cut-off score for 'pain'(positive), a parameter for false negative scores. Specificity: the percentage of control horses that is also scored below the cut-off score for 'no pain'(negative), a parameter for false positive scores. The negative predictive value: the percentage of horses that are scored below the cut-off score for 'no pain' that are really negative (controls). The positive predictive value: the percentage of horses that are scored above the cut-off score for 'pain' that are really positive(patient /colic horses).

Within the colic patient group (22 horses at *T= (0)*) patients were either treated conservative (presumably corresponding with 'moderate pain') or had surgery or euthanasia and corresponded presumably with 'severe pain'. The CPS appeared to be sensitive and specific to distinguish 'severe pain' from 'moderate pain'. This resulted in a sensitivity of 87,5% and a specificity of 85%.

Specificity of separate elements of CPS

Figure 12 shows the frequency of scores per separate element of the CPS for all horses scored. The frequency of each element scored is compared between patient and control horses. Four elements (rectal temperature, laying down/rolling, sweating and kicking at abdomen) never received a score other than 0 (normal) in control horses. Digestive sounds and posture scored much higher in colic horses compared to controls, based on face value. (Tabe 6) The remaining elements were scored higher in colic horses compared to control horses, except for 'Tail flicking' and 'head movements'.

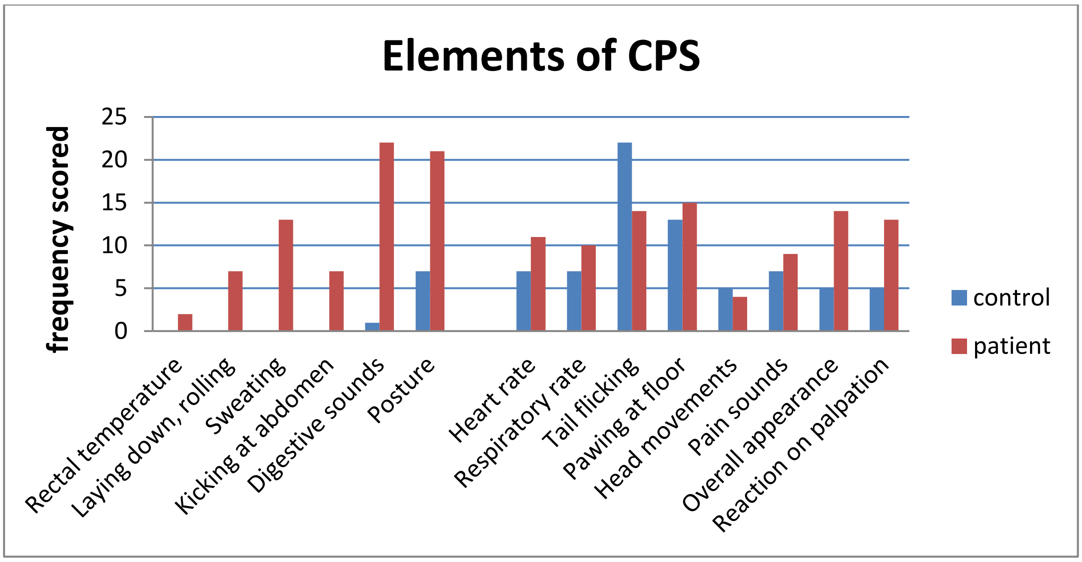

figure 12
Separate elements of CPS; red bars represent the frequency of scoring that element of all patient horses. Blue bars represent the frequency of scoring the element for control horses. The six elements on the left are those elements,which either scored only in the patients and never in the controls. Or were at least scored 3x more in frequency for patient horses on that element than for the control horses. Based on face value.

Z-value of separate elements

	Rectal temperature	Laying down, rolling	Sweating	Kicking at abdomen	Digestive sounds	Posture	Heart rate	Respiratory rate	Tail flicking	Pawing at floor	Head movements	Pain sounds	Overall appearance	Reaction on palpation
patient	2	7	13	7	22	21	11	10	14	15	4	9	14	13
control	0	0	0	0	1	7	7	7	22	13	5	7	5	5
z-value	1	1	1	1	0,91	0,5	0,21	0,18	-0,22	0,07	-0,11	0,13	0,47	0,44

Table 6
The scoring frequency of each element is shown for patient and control horses. Z- value is calculated by (patient-control)/(patient+control).

CPS timeline

All patients that have been scored more than once by the CPS are shown in table 7. The median and standard deviation of all patients of each measured moment are shown in figure 13a. The scores on *T= (1st morning)* and *T = (2nd morning)* were lower than the scores in the acute phase *T= (0), T = (1), T = (2-3).* Figure 13b shows the CPS scores of all horses separately.

Scores on CPS – median & standard deviation

Patient code		0	1	2,5	24	48
K10	1	14,5			2,5	5
K12	2	5,5			2,5	
K13	3	10			1	3,5
K17	4	5,5			2	0,5
K18	5	6		1,5	5,5	6,5
K19	6				7	6
K20	7			6	5	3
K21	8		8		7	11,5
K23	9	17	15		6,5	
K24	10	10,5			8	
K25	11	6		8,5		
median		8	11,5	6	5,25	5
Std dev.		3,971	4,950	3,547	2,497	3,460

h after arrival	Measure moment
0	T0
1	T1
2.5	T2-3
24	T1st morning
48	T2nd morning

Table 7
Mean scores of the two observers on CPS of all horses that are measured at different times after arrival (N=11). K= ID number of colic horse in this study.

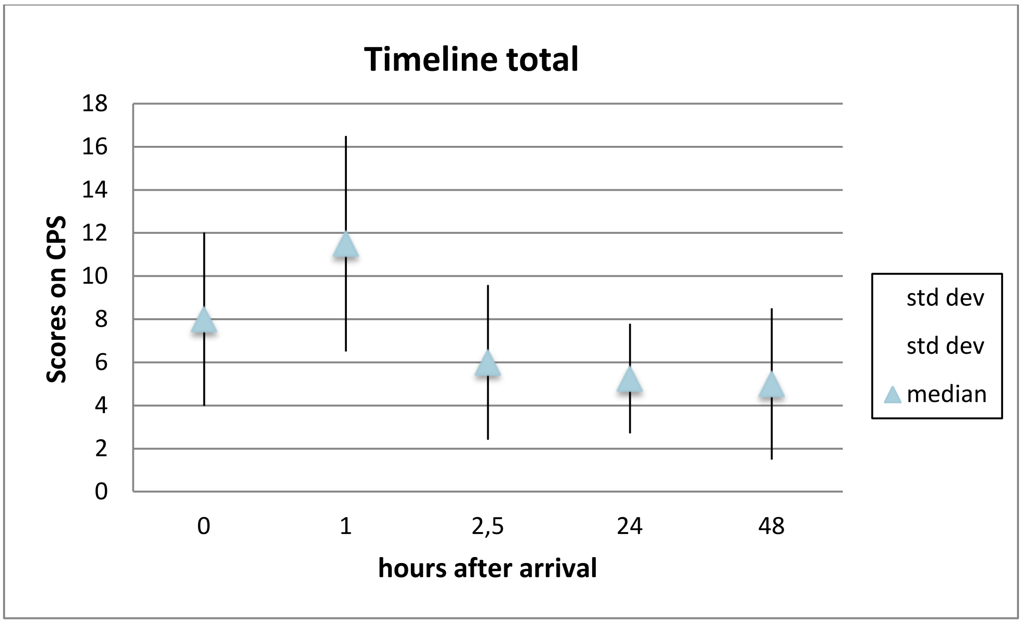

Figure 13a
The blue arrow points are the median scores on CPS of all horses scored at that moment, the black lines show the standard deviation matching the blue arrow point of that moment.

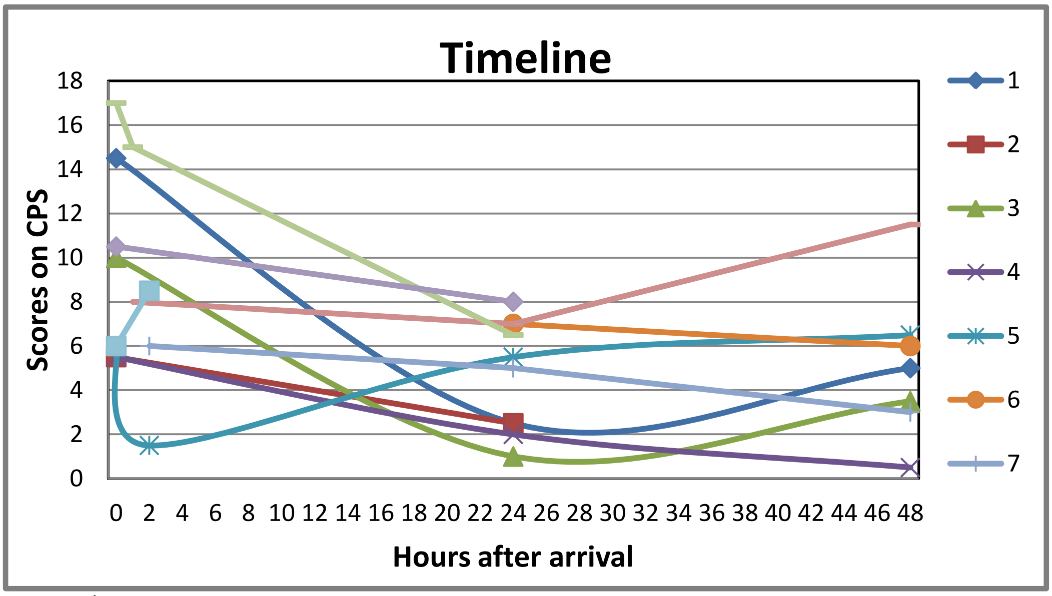

Figure 13b
Timeline of CPS scores on different times after arrival of every horse (table 6) separately.

Personality

The questionnaires of 13 patients were used for determining personality (table 8). PCA with Varimax rotation on the data collected from the questionnaires extracted 2 components accounting for 84,9% of the total variance. The first component, accounting for 56,9% of the total variance, loaded strongly (higher than +0,6 or lower than -0,6 (Ijichi et al., 2013)) positive for Neuroticism and Extroversion and negative for Gregariousness to people. Because these traits were shown to be related to "pro-active coping" this component was labeled 'proactive'. The other 28,0% of the total variance was included in the second component. Agreeableness and Gregariousness to horses loaded strongly positive to this component. These two personality traits fitted with a reactive coping style, so this component was labeled 'reactive'. All is shown in table 8 and figure 14.

Loading of different factors to the extracted components

	Component	
	1 - "proactive"	2 - "reactive"
Agreeableness (A)	-0,217	**0,945**
Neuroticism (N)	**0,714**	-0,294
Extroversion (E)	**0,842**	-0,339
Greg. people (F4)	**-0,957**	-0,094
Greg. Horses (F5)	-0,109	**0,973**

Table 8

Component 1 - 'proactive'; Neuroticism and Extroversion are loading positively and Gregariousness to people is loading negatively to this component.

Component 2 - 'reactive'; Agreeableness and Gregariousness to horses are loading positively to this component.

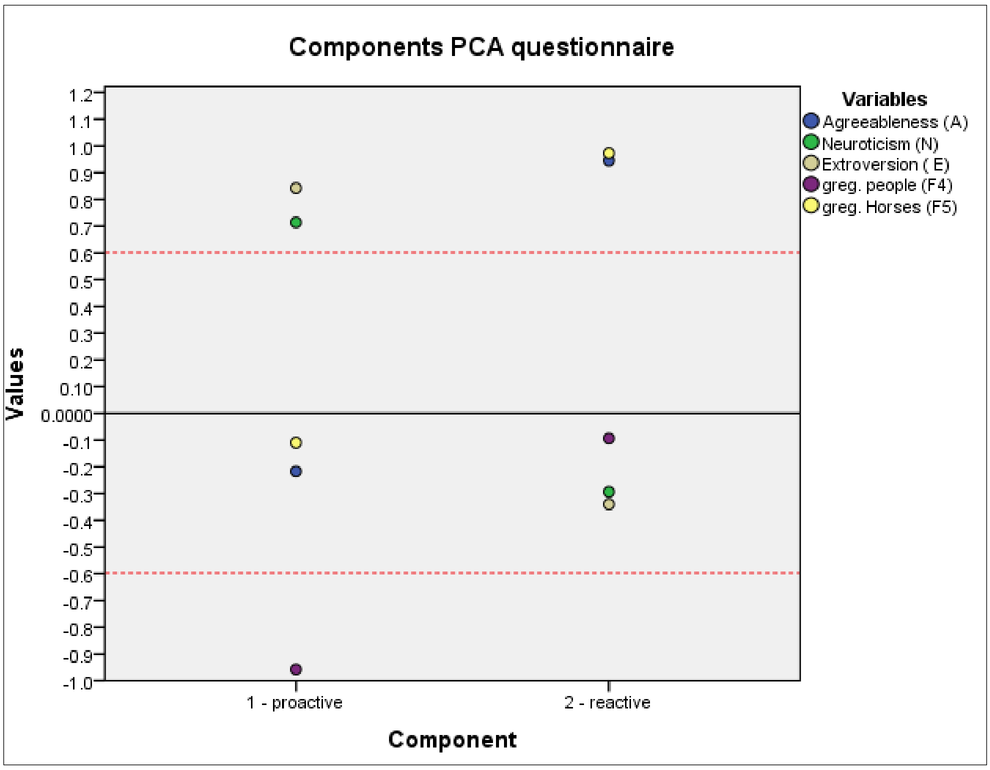

Figure 14
Overview of loading of all factors for each component. Red dotted line is the point where above (+0,6) or beneath (-0,6) the factor is loaded strongly to the component.

Overview of scores on Questionnaire & CPS and grouping to component 1 or 2

	Mean scores questionnaire					CPS scores		Personality
	A	N	E	F4	F5	T0 mean	T0 cons	Component
K10	5,0	1,3	1,4	5,0		14,5	14,5	
K12	4,1	2,1	2,8	5,0	5,0	5,5	5,5	2 – reactive
K13	3,9	1,9	2,7	4,0	4,0	10,0	10,0	1 – proactive
K15	4,7	1,0	1,8	5,0	5,0	12,0		2 – reactive
K17	4,6	1,0	2,2	5,0		5,5	5,5	
K18	4,6	1,2	2,2	5,0	5,0	6,0	6,0	2 – reactive
K20	4,3	2,4	3,1	4,0	4,0			1 – proactive
K21	1,8	2,8	4,1	4,0	2,0			1 – proactive
K23	4,4	3,2	4,8	1,0	5,0	17,0	17,0	1 – proactive
K24	5,0	1,6	3,0	4,0	5,0	10,5	10,5	2 – reactive
K25	4,4	2,6	3,0	5,0	4,0	6,0		1 – proactive
K26	4,8	1,4	2,8	5,0	4,5	10,5	10,5	2 – reactive
H08	5,0	2,8	1,0	5,0	5,0			2 – reactive

Table 9
K= ID number of colic horse in this study that had at least one completed questionnaire, H= patient with pain in the head also with a completed questionnaire. Scores on questionnaires were scores for A: agreeableness, N: neuroticism, E: extroversion, F4: gregariousness to people, F5: gregariousness to horses. Green cells: mean of two completed questionnaires. Missing scores for F5 (K10 and K17) are due to not fully completed questionnaires. CPS scores: mean scores of two observers (at T=(0)). Blue rows show the patients that were usable for comparing the CPS scores between 'reactive' and 'proactive' horses (N= 8), rows without color were missing either a score on the questionnaire or at *T=0* on CPS (not usable). Dark blue rows show the patients that are treated conservatively (*T=0 cons)* (N=6). Personality: the division of the patients between the 2 extracted components 'proactive' or 'reactive'.

For every patient with a complete dataset (all questions answered) (N = 11) the correlation between the scores on all factors and the extracted components was determined (Table 9). All horses with green marked cells did have two completed questionnaires. The mean scores of the two questionnaires were used. Two patients (K10 and K17)did have just one questionnaire which was not fully completed, the last 3 questions were missing (no F5 data). For the determination of the personality components these two horses were excluded. Every patient was classified to the component with the highest positive correlation or the least negative correlation.
The mean CPS pain scores on *T= (0)* (8 patients labeled blue in table 9) were compared between 'proactive' and 'reactive' horses. There was little difference comparing the two with the box plot (figure 15) and Mann Whitney U test showed a significance of P=0,786 which means that there is no significant difference (for P<0,05). Comparing just the 6 conservatively treated patients (labeled dark blue in table 9) the box plot (figure 16) was not significant either (P = 0,533).

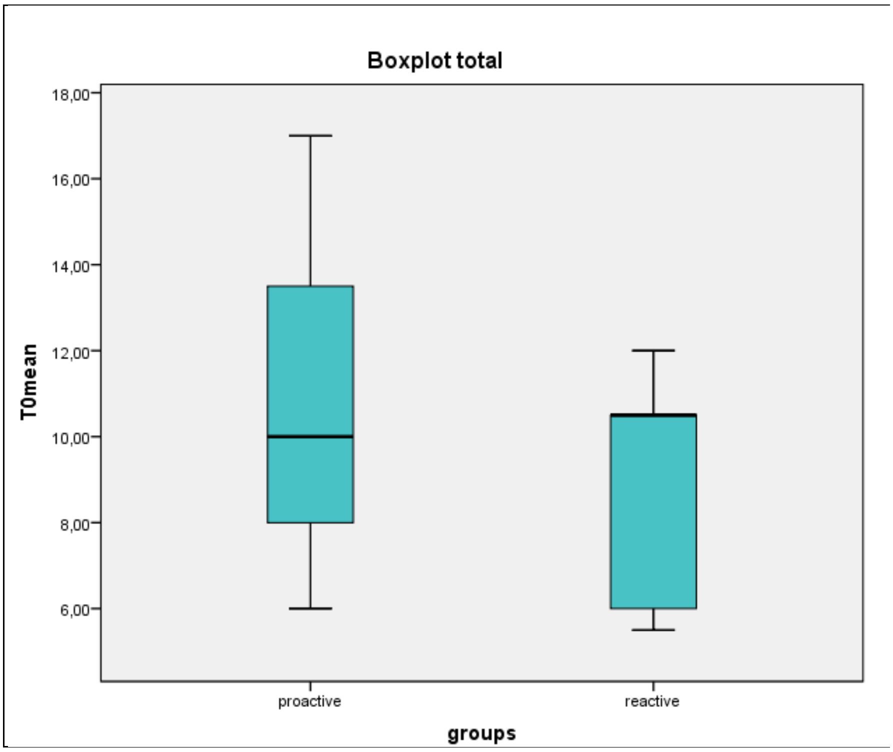

Figure 15
CPS scores of all horses at T = (0) are compared between proactive and reactive labeled horsed, P=0,786. N=8, 5 reactive, 3 proactive. Box plots show the spread of all scores in that component, blue box: median 50% of scores, black line: median score.

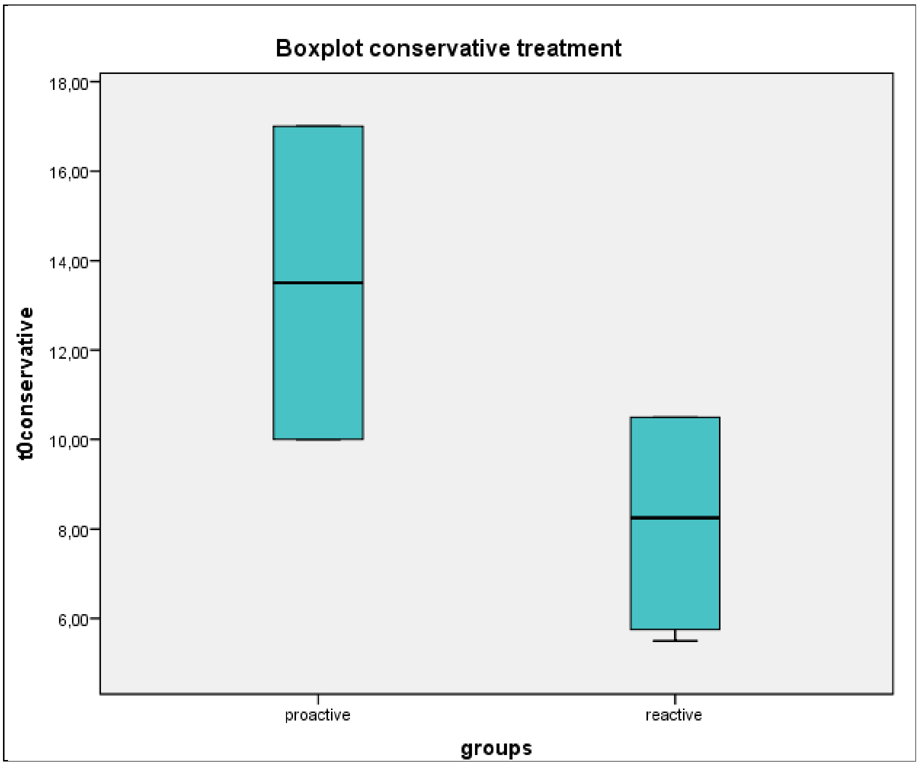

Figure 16
CPS scores of all conservative treated horses at T = (0) are compared between proactive and reactive labeled horsed, P=0,533. N=6, 4 reactive, 2 proactive. Box plots show the spread of all scores in that component, blue box: median 50% of scores, black line: median score.

Discussion

Pain scales

Wagner (2010) described: 'Potential advantages of using a defined pain scoring system or pain scale are:
(1) enhancement of the observer's ability to recognize and potentially quantify pain in individual animals,
(2) enhancement of inter-observer agreement on the degree of pain assessed, improving patient care with more consistent evaluations,
(3) tendency to raise awareness of animal pain and encourage treatment, especially when the scale itself designates a certain minimum score above which treatment must be instituted
(4) ability to follow and document progress or deterioration in patient condition through decreases or increases in pain score'(Wagner, 2010)

The results of this study will be discussed on the basis of this description. The green numbers matching to the points described by Wagner (2010) discuss the positive outcomes of this study, the red numbers below discuss the factors in this study that could be improved.

(1) The CPS was designed to be useful for untrained observers to objectively score pain in horses having colic. Observers were able to score all elements within 5 minutes at a horse with colic. The observers in this study did have just read the scale a few times before starting to score the patients. There was no discussion about the elements when scoring. In this study the observers were veterinary students who had already learned to measure physiologic parameters like heart rate and digestive sounds. It is likely that experienced veterinarians are able to score the parameters without training as well.
No learning curve in rating CPS scores was demonstrated in this study (figure 9). Differences in scoring between the two observers were mainly in the high overall scores. So the differences could be explained that by the fact that severe colic patients showed a 'high' pain score with more variation in scores compared to the controls or moderate colic patients with a 'mild' pain score. There was no larger difference in the beginning than at the end of the period of data collecting. The scores of 3 and 4 October were from control horses only, on these dates there were low or no differences in scoring between the observers.

(2) Inter-observer reliability was high and significant at a 0,001 level for CPS and FEPS. Both scales could be useful to discriminate significantly (p <0,001) between patient and control group (= sensitivity). On the other hand the VAS as 'current standard' showed a lower inter-observer correlation, however, still significant at a 0,05 level (fig 8). Other studies after CPS showed a high inter-observer reliability (Bussieres et al., 2008;Lindegaard et al., 2010; van Loon et al., 2010; Van Loon et al., 2014). In the study of Lindegaard et al. (2010) VAS inter-observer reliability was classified as fair (Lindegaard et al., 2010). In this study the VAS did

not differentiate between patient and control group (table 5). However, in studies in human emergency departments, where patients rated their own pain acute severity, this VAS showed to be sufficiently reliable (Bijur et al., 2001). On the other hand, even though a VAS seemed reliable and valid it was not usable for owners to rate the pain of their dog which had osteoarthritis, since owners were not able to recognize the pain of the animal)(Hielm-Björkman et al., 2011).This study showed that VAS was not always an objective scale to be used by humans (in this case veterinarians) to rate the pain of an animal.

(3) Sensitivity and specificity taking into account, our analysis demonstrated that a VAS was less reliable than CPS, FEPS or CPS+ FEPS combined. CPS, FEPS and CPS+FEPS each had a high sensitivity and specificity to distinguish between colic horses and controls. CPS+FEPS combined gave the best outcome in sensitivity and specificity (table 5). This could be explained by that the combination of more different elements (facial combined with total) in scoring pain resulted in an even more reliable pain assessment. Even though FEPS was designed to score pain in general and not specifically abdominal pain, the CPS could be improved by including some facial parameters to score pain based on more objective elements.

Within the colic horses the CPS showed a high sensitivity and specificity to distinguish between 'severe colic' (resulting in surgery or euthanasia) and 'moderate or mild colic' (treated conservatively). This could be useful for veterinarians in the clinic for the estimate whether surgery or euthanasia (or severe pain killers) will be needed for a colic patient. The VAS scores did not show a sufficient sensitivity and specificity to distinguish the patients from the controls. Therefore with a VAS no differentiation could be made between 'mild / moderate' or 'severe' pain due to colic in horses. This in combination with a lower inter-observer reliability than CPS and FEPS showed that VAS seems no longer the best pain assessment scale for horses with acute colic.

(4) In this study a follow up of the development of the pain was done by scoring each horse on different times. The goal was to monitor the acute phase (*T=(0)/T= (1)/T= (2-3)*) and the two days afterwards (*T= (1st morning), T = (2nd morning)*). In practice it turned out to be difficult to monitor the horses on all those different times. Table 7 and figures 13a and 13b show the results of the horses that had been scored more than once using CPS. Most patients showed a decrease on CPS score after 24 hours (and 48 hours as well) when receiving a conservative treatment (figure 13a). Increased scores could be explained by different factors. For example: Patient nr 11 in figure 13b showed an increase during the acute phase (the first hours after arrival). This horse was later euthanized due to increasing pain due to severe colic (pain killers did no longer work). The horses that showed an increase after a dip in the scores were likely to express stress behavior rather than pain. In our experience these were horses that responded stressfully on a new environment when moved to the colic box for filming. (young horses, extravert horses, a mare without foal). The influence of personality of horses will be discussed later.

The CPS now seems to be a more reliable pain scoring system. There are however factors in the experimental design that could be improved to acquire a more objective and reliable outcome of this study. This will be discussed based on the points of Wagner (2010) in the next paragraphs.

(1) The descriptions of the elements seemed to be well defined. Still it would be less discussable if some elements were clarified with pictures/photos or video of a horse showing that behavior (head movement, posture, overall appearance).
Veterinarians and veterinary students were able to use the scale without training. When the scale would be also available for general horse owners as well, they cannot measure the physiological parameters reliably. The behavioral elements seemed to be well described and when pictures/photos or video could be added it can be measured in a new study whether owners can score those elements reliably.

Four elements (rectal temperature, laying down/rolling, sweating and kicking at abdomen,) (figure 12) did never get a score above 0 (normality) in control horses ($z = 1$). In this group of horses, these four elements were specific for horses with acute colic, so scoring these elements indicated a horse having pain from colic. However this does not mean that the other way around not scoring on these elements indicated a control/pain free horse.
Digestive sounds and posture scored much higher in colic horses compared to controls (respectively $z = 0{,}91$ and $z = 0{,}5$).
All other remaining elements were scored higher in colic horses compared to control horses, except for 'tail flicking' ($z = -0.22$) and head movements ($z = -0.11$). This meant that these remaining factors were possibly associated with pain due to acute colic, but also could be expressed due to another stress factor. Stress due to something else than pain (new environment, social isolation etc) could therefore be a factor that influenced the score on these elements (Taylor et al., 2002).
In the current CPS all elements were equal in weight. To obtain a more specific scale for colic pain the (five) elements specific for colic could obtain higher weighting factors than the elements that score associated with general restlessness. Some elements (like 'tail flicking' and head movements) that did not differentiate between colic and control horses could be removed from the scale or should be described more specifically for signs of colic (for example making combinations: teeth grinding just counts if the score on posture is ≥2).
Other studies found different elements specific for pain. Bussieres et al. (2008) described 'posture', 'kicking at abdomen', 'interactive behaviour' and 'palpation of painful area' as good to excellent specificity for the pain measured by CPS. Different from this study was that the study of Bussieres et al. (2008) was measuring (induced) orthopedic pain (Bussieres et al., 2008). The study of Van Loon et al. (2014) after horses with visceral pain showed that

pawing on the floor, appearance, head movements and interactive behavior were the most correlating elements with the total CPS score (Van Loon et al., 2014).
In both studies 'sweating' was a non-sensitive parameter for pain (Bussieres et al., 2008; Van Loon et al., 2014), while this study showed 'sweating' as one of the four most specific parameters for acute colic in horses.

(2) Currently there is no validated 'golden standard' for validation of CPS (Van Loon et al., 2014). In this study VAS was used like in the study of Lindegaard et al. (2010). VAS is the most comparable scoring system with the current pain scoring method in horses: subjective scoring on experience of the veterinarian. In human medicine VAS is proved reliable for rating the pain of patients (Bijur et al., 2001). In the study of Lindegaard et al. (2010) induced orthopedic pain was scored by means of a CMPS (composite measure pain scale) and by means of a VAS. There was a fair agreement between the two scoring systems, de regression line between the scoring systems did not deviate significantly from the line of equality (Lindegaard et al., 2010). Different from this study was that the CMPS contained subjective elements and there were no physiological parameters in the scoring system. The VAS was scored by the same observers as the CMPS. This could explain the difference in outcome of our study and the study of Lindegaard et al. (2010).
Van Loon et al. (2014) used the NRS (Pritchett et al., 2003) for comparing to CPS. In this study was decribed that the NRS was not suitable for validating the CPS because of the lack of validation of the NRS and the moderate inter-observer reliability (Van Loon et al., 2014). The NRS is designed for scoring pain in colic horses, but contains subjective elements and no physiologic parameters (Pritchett et al., 2003). This is why CPS is thought to be more reliable.
Graubner et al. (2011) used the differences in pain scores between horses with painful complications an horses without complications to validate the PASPAS (post abdominal surgery pain assessment scale) because of the lack of a 'golden standard'. This is a composite pain scale that contains both physiologic and behavioral parameters, but also contains a subjective element 'general pain signs'. This scale had the tendency to be valid and reliable to score pain after abdominal surgery (Graubner et al., 2011).
So there is still no possibility to validate the CPS to another objective composite numerical rating scale (containing both physiologic and behavioral parameters). A numerical rating scale like PASPAS or NRS seems to be more applicable for validation of CPS than VAS.

The observers of CPS and FEPS in this and the previous study were not blinded for diagnosis, therapy and patient/control groups. This could influence the outcome of the CPS scores. The one-minute film clips to score the VAS on the other hand were blinded and randomized in patient or control horses. An option to improve is to just film the colic horses for 5 minutes and afterwards randomize the patient and control horses. Veterinarians could be asked to score VAS based on the first minute of each film clip. Same veterinarians could be asked to score the randomized 5 minute film clips on CPS. It was not performed this way in this study, because of practical reasons; it would take twice as much time to get the results (film

and score apart).
For scoring the VAS the veterinarians did have just the first minute of the 5 minutes filmed. Some patients did not express any pain in that first minute or controls seemed like to have pain, because of the expression of stress induced behavior. In our experience this stress induced behavior was mainly showed in the first 1-2 minutes in the colic box. After that it was easier to differentiate between colic and control horses. So showing the veterinarians the whole 5 minutes, or 1 minute in the middle or end of the 5 minutes filmed could change their opinion of the pain intensity of the horse.
Because of technical problems, not all patients scored on CPS and FEPS were recorded on film. This resulted in less film clips for VAS to be scored by veterinarians and could have contributed to the lower inter-observer reliability and sensitivity/specificity compared to the CPS and FEPS scores.

(3) Sensitivity and specificity of VAS, CPS, FEPS and combined CPS+FEPS scores were based on cut-off scores set on experience. This is still a subjective way of determining when CPS scores are high enough to count as 'pain'. This was the first time these cut-off scores were used. CPS appeared to be able to distinguish controls and colic horses and to differentiate 'mild colic' from 'severe colic' based on the cut-off scores described in the material and method section (pg 13).
Van Loon et al. (2014) described that it is still challenging to determine objective cut-off scores for (nonverbal) animals (Van Loon et al., 2014). It would be beneficial if there will be objective cut-off scores to differentiate between 'no pain', 'mild pain' and 'severe pain' as a basis for the treatment and prognosis.
Bussières et al. (2008) determined sensitivity based on the possibility of the scale to distinguish between three 'pain groups' (E-groups). Specificity was based on de percentage of presence of pain scores in the 'control groups' (C- groups). Based on these parameters the sensitivity of total CPS was good-to-excellent and the specificity was good (Bussieres et al., 2008).

(4) A timeline (figure 13) of the CPS scores could be a good way to follow and document progress or deterioration in patient condition. However, the timeline in this study was not yet reliable because there were too few horses, which could be followed for several time points. Best would be to monitor each horse at all times (*T= (0), T=(1)/T=(2-3), T= (1st morning), T= (2nd morning)*). Especially in the acute phase this could give information about the influence of pain medication on the colic pain. *T= (1) h* should show a decrease in pain score, when given NSAID's at the clinic on *T= (0)*. *T= (2-3) h* could show an increase in pain score compared to *T= (0),* when the pain medication was given at home or was unknown, because the pain medication could no longer be working. Differences in the acute phase (T = (0), T = (1) and T = (2-3) were not yet possible to show in this timeline.

It would be beneficial to collect more data for a greater reliability of these results. The results in this article are based on all data collected including this study and the previous study. Statistical analysis on just either one separately resulted in a non-significant outcome, but combining the two resulted in that the significance improved to multiple significant outcomes.

Personality

Questionnaire

To get a reliable outcome of the questionnaires raters should be able to score objectively. Best ratings would be obtained by responding the questionnaires by trainers/owners of the horse that know the horse for at least 6 months and handle the horse for at least 4 times a week (Lloyd et al., 2007). Even better would be if at least two handlers could complete the questionnaire. This criteria could not be strictly followed in this study. The person who accompanied the horse to the clinic and spent the most time with the horse was asked to fill in the questionnaire. In most cases one handler completed the questionnaire. The questionnaire was also filled in by a student, which had treated the patient for a few days. However, it is questionable if this gave valuable information: the horse was 'ill', was being housed in a strange environment and was receiving different (sometimes painful, when receiving IM or IV injections or being rectally explored) treatment compared to their normal life. This could influence the interpretation of their behavior (Taylor et al., 2002). Besides that it is hard to know a horse really well in a few days working with the horse a few minutes a day.

In this study the inter-observer reliability of the used questionnaire could not be tested, but the study of Lloyd et al. (2007) demonstrated high levels of agreement between raters using a comparable questionnaire (Lloyd et al., 2007). Based on these findings we decide that all personality data collected in this study could be used in the PCA calculations. In this study we chose to use all the collected questionnaires to obtain the most information. Normally a PCA gives the best results when twice as many individuals compared to the number of variables can be included (Lloyd et al., 2007) In this case 13 horses and 5 variables were included. Generally spoken, this number was too small to obtain a reliable outcome. Figure 16 shows there was a difference (not significant P 0,553) in CPS scores, so testing this again using more data could result in a statistically significant outcome.

A goal for a next study would be to acquire more completed questionnaires per individual preferably completed by handlers that know the horse for a long time and to test more individuals during the study.

Behavioral tests

The performed behavioral tests were not used for analysis in the PCA. That would have resulted in a too big number of variables resulting in a less reliable outcome of the PCA. In addition the behavioral tests showed little differences between horses. This would therefore not have been a contribution to the distinction of personalities.

However, there were some elements in the testing procedure that could be improved to acquire more reliable results: the testing was performed on *T= 2nd morning.* Patients with colic, which were treated conservatively normally receive only a little bit of food on the second day at the clinic after few days fasting. This could influence the outcome of especially the 'garbage bag test'(Christensen et al., 2008), because they could be more focused on food and thus show less reaction on the removing on the garbage bag, compared to when fed normal amounts of food. This means that a difference between colic horses and control horses was expected in the reaction on the 'garbage bag test'.
Environmental influences should be minimized because this could change reaction on several tests. In this study people, material, and test area (colic box) were standardized, but there were some differences in environment like noise in the corridor in front of the colic box, people/horses in the box next to the colic box or noise behind the doors to outside. Horses could be focused on that noise and therefore react less to people, sound or material in the test. The standard whistle tone of the mobile phone ('negative sound' test) for example seemed not loud enough to provoke a reaction when there was environmental noise. Using a loud whistle used in sports instead of the standard whistle tone of the mobile phone could improve the outcome of this part of the tests.
More tests should be performed to determine if the behavioral tests contributed to the assessment of personality of the horse.

Influence of personality on CPS score

The influence of personality of the horses with colic on a CPS score was tested for the first time with this study. It was expected that extravert/proactive animals could show more stress behaviour when feeling pain compared to introvert/reactive animals, what could result in a higher (behavioural) score on CPS. In the article of Ijichi et al.(2013) horses were scored by their owners by means of the same personality questionnaire as used in this study. According to Ijichi et al. (2013) horses that scored high on Neuroticism seemed to have a reduced threshold for pain (lameness) and horses that scored high on Extroversion expressed pain more overtly independent of the severity of pain (Ijichi, Unpublished, Draft version. 2013). In this study also both Neuroticism and Extroversion were strongly positively loaded to 'proactive' personality of horses (table 8 + figure 14). However, we could not yet collect enough data were collected to get a significant outcome. Figures 15 and 16 show there was a tendency of a difference between proactive and reactive horses in CPS scores, where proactive horses scored higher (as expected).
Because this was the first time the influence of personality on CPS was examined the study design was broad. Mean scores on *T= (0)* for all the 8 blue labelled horses (table 9) and for the 6 conservatively treated horses (labelled dark blue in table 9) were divided based on 'proactive' or 'reactive' personality. Subsequently the CPS scores of the "proactive" group were compared to the CPS of the "reactive" group. The comparison within all 8 horses did not show a significant outcome (P=0,786 figure 15), neither within just the 6 conservative treated horses (P=0,533 figure 16)

More specific should have been to compare pain scores of patients having the same type of colic (left ventral colon constipation (LVC), right dorsal displacement (RDD), strangulation etc.), because this would partly exclude the variation in severity of pain within the groups. That could be a reason why the comparison within the conservative treated horses tends to be more significant than the comparison within all the colic patient.
In this study horses were strictly classified in either ‘proactive’ or ‘reactive’ to make groups comparable. Better would be to use just the horses that loaded strongly to one of the two components instead of using all horses even if most of them were in between ‘proactive’ and ‘reactive’.
To make these adjustments more data should be collected. Some more (specific) tests should be performed to possibly validate the hypothesis that there is influence of personality on the CPS scores in colic horses.

Conclusion

This study showed a high inter-observer reliability and a high sensitivity and specificity for pain assessment by means of CPS in this group of patients. CPS could discriminate between horses with ‘mild/moderate’ (conservative treated) and ‘severe’ colic (surgery/euthanasia). The ‘current standard’ VAS could not significantly differentiate between colic and control horses (low sensitivity and specificity) and showed a low inter-observer reliability although significant. There was a weak correlation between CPS and VAS. From this could be concluded that CPS was a better tool to assess pain due to acute colic compared to VAS. Although some adjustments could improve the scale, CPS seemed a useful tool for assessment of abdominal pain in horses.

The use of questionnaires and behavioral test on personality has not yet provided sufficient data to significantly show a difference between the CPS scores for ‘proactive’ and ‘reactive’ horses. There seemed to be a tendency in ‘proactive’ horses for having a higher CPS, but more research should be done after the influence of personality of the horse on the pain scored using CPS.

References

Ashley, F.H., Waterman-Pearson, A.E., Whay, H.R., 2005. Behavioural assessment of pain in horses and donkeys: application to clinical practice and future studies. Equine Vet. J. 37, 565-575.

Bijur, P.E., Latimer, C.T., Gallagher, E.J., 2003. Validation of a verbally administered numerical rating scale of acute pain for use in the emergency department. Acad. Emerg. Med. 10, 390-392.

Bijur, P.E., Silver, W., Gallagher, E.J., 2001. Reliability of the visual analog scale for measurement of acute pain. Acad. Emerg. Med. 8, 1153-1157.

Bussieres, G., Jacques, C., Lainay, O., Beauchamp, G., Leblond, A., Cadore, J.L., Desmaizieres, L.M., Cuvelliez, S.G., Troncy, E., 2008. Development of a composite orthopaedic pain scale in horses. Res. Vet. Sci. 85, 294-306.

Christensen, J.W., Malmkvist, J., Nielsen, B.L., Keeling, L.J., 2008. Effects of a calm companion on fear reactions in naïve test horses. Equine Vet. J. 40, 46-50.

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

Graubner, C., Gerber, V., Doherr, M., Spadavecchia, C., 2011. Clinical application and reliability of a post abdominal surgery pain assessment scale (PASPAS) in horses. Vet. J. 188, 178-183.

Hicks, C.L., von Baeyer, C.L., Spafford, P.A., van Korlaar, I., Goodenough, B., 2001. The Faces Pain Scale – Revised: toward a common metric in pediatric pain measurement. Pain 93, 173-183.

Hielm-Björkman, A.K., Kapatkin, A.S., Rita, H.J., 2011. Reliability and validity of a visual analogue scale used by owners to measure chronic pain attributable to osteoarthritis in their dogs. Am. J. Vet. Res. 72, 601-607.

Ijichi, C.e.a., Unpublished, Draft version. 2013. Pain expression is linked to personality in horses. .

Ijichi, C., Collins, L.M., Creighton, E., Elwood, R.W., 2013. Harnessing the power of personality assessment: subjective assessment predicts behaviour in horses. Behav. Processes 96, 47-52.

Koolhaas, J.M., Korte, S.M., De Boer, S.F., Van Der Vegt, B.J., Van Reenen, C.G., Hopster, H., De Jong, I.C., Ruis, M.A.W., Blokhuis, H.J., 1999. Coping styles in animals: current status in behavior and stress-physiology. Neuroscience & Biobehavioral Reviews 23, 925-935.

Langford, D.J., Bailey, A.L., Chanda, M.L., Clarke, S.E., Drummond, T.E., Echols, S., Glick, S., Ingrao, J., Klassen-Ross, T., Lacroix-Fralish, M.L., Matsumiya, L., Sorge, R.E., Sotocinal, S.G., Tabaka, J.M., Wong, D., Van Den Maagdenberg, A.M.J.M., Ferrari, M.D., Craig, K.D., Mogil, J.S., 2010. Coding of facial expressions of pain in the laboratory mouse. Nature Methods 7, 447-449.

Lindegaard, C., Thomsen, M.H., Larsen, S., Andersen, P.H., 2010. Analgesic efficacy of intra-articular morphine in experimentally induced radiocarpal synovitis in horses. Veterinary Anaesthesia and Analgesia 37, 171-185.

Lloyd, A.S., Martin, J.E., Bornett-Gauci, H.L.I., Wilkinson, R.G., 2008. Horse personality: Variation between breeds. Appl. Anim. Behav. Sci. 112, 369-383.

Lloyd, A.S., Martin, J.E., Bornett-Gauci, H.L.I., Wilkinson, R.G., 2007. Evaluation of a novel method of horse personality assessment: Rater-agreement and links to behaviour. Appl. Anim. Behav. Sci. 105, 205-222.

Love, E., Gillespie, L., Colborne, G., Conference abstract. 2011. Facial expression of pain in horses. .

Mills, D.S., 1998. Personality and individual differences in the horse, their significance, use and measurement. Equine Vet. J. Suppl. (27), 10-13.

Minero, M., Dalla Costa, E., Lebelt, D., Stucke, D., Canali, E., Leach, M., 2013. Development of a facial expressions pain scale in horses undergoing routine castration.

Molony, V., Kent, J.E., 1997. Assessment of Acute Pain in Farm Animals Using Behavioral and Physiological Measurements. J. Anim. Sci. 75, 266-272.

Momozawa, Y., Ono, T., Sato, F., Kikusui, T., Takeuchi, Y., Mori, Y., Kusunose, R., 2003. Assessment of equine temperament by a questionnaire survey to caretakers and evaluation of its reliability by simultaneous behavior test. Appl. Anim. Behav. Sci. 84, 127-138.

Muir, W.W., 2010. Pain: Mechanisms and Management in Horses. Veterinary Clinics of North America - Equine Practice 26, 467-480.

Price, J., Catriona, S., Welsh, E.M., Waran, N.K., 2003. Preliminary evaluation of a behaviour-based system for assessment of post-operative pain in horses following arthroscopic surgery. Vet. Anaesth. Analg. 30, 124-137.

Pritchett, L.C., Ulibarri, C., Roberts, M.C., Schneider, R.K., Sellon, D.C., 2003. Identification of potential physiological and behavioral indicators of postoperative pain in horses after exploratory celiotomy for colic. Appl. Anim. Behav. Sci. 80, 31-43.

Sutton, G.A., Dahan, R., Turner, D., Paltiel, O., 2013. A behaviour-based pain scale for horses with acute colic: Scale construction. The Veterinary Journal 196, 394-401.

Taylor, P., Pascoe, P., Mama, K., 2002. Diagnosing and treating pain in the horse. Where are we today? The Veterinary clinics of North America.Equine practice 18, 1-19, v.

Unknown, 2012. Welzijnsmonitor Paardenhouderij – het meetprotocol – versie 3.0. report nr 572.

van Loon, J.P.A.M., Back, W., Hellebrekers, L.J., van Weeren, P.R., 2010. Application of a composite pain scale to objectively monitor horses with somatic and visceral pain under hospital conditions. Journal of Equine Veterinary Science 30, 641-649.

Van Loon, J.P.A.M., Jonckheer-Sheehy, V.S.M., Back, W., René van Weeren, P., Hellebrekers, L.J., 2014. Monitoring equine visceral pain with a composite pain scale score and correlation with survival after emergency gastrointestinal surgery. Veterinary Journal 200, 109-115.

Wagner, A.E., 2010. Effects of Stress on Pain in Horses and Incorporating Pain Scales for Equine Practice. Veterinary Clinics of North America - Equine Practice 26, 481-492.

Annex

Annex I - CPS

Date:

Time:

Observer:

Observation:

T0 hours

T1 hour — if given NSAIDs at clinic

T2 – 3 hours — if pain management is unknown or was administered before arrival at clinic → optional

T1st morning

T2nd morning

Data	Categories	Score
Physiological data Heart rate	24 - 44 beats/min 45 - 52 beats/min 53 - 60 beats/min > 60 beats/min	0 1 2 3
Respiratory rate	8 - 13 breaths/min 14 - 16 breaths/min 17 - 18 breaths/min > 18 breaths/min	0 1 2 3
Rectal temperature	36.9 °C - 38.5 °C 36.4 °C - 36.9 °C or 38.5 °C - 39.0 °C 35.9 °C - 36.4 °C or 39.0 °C - 39.5 °C 35.4 °C - 35.9 °C or 39.5 °C - 40.0 °C	0 1 2 3
Digestive sounds	Normal motility Decreased motility No motility Hypermotility or steelband	0 1 2 3
Behavior Posture	Quietly standing and/or one hind leg resting, explores environment Slightly tucked up abdomen, still explores environment (with possible unrest) Extremely tucked up abdomen, hunched back and/or stretching of body/limbs Does not stand or for short amounts of time (<1 min), sits on hindquarters	0 1 2 3

Laying down, rolling	Does not lie down or rests lying down Lies down in normal posture, rolls or tries to roll (1-2 times/5 min) Alternates lying down and standing, rolls or tries to roll (>2 times/5 min) Constantly lies in an abnormal position: on its side with stretched limbs, on its back, or does not stop rolling	0 1 2 3
Sweating	No signs of sweating Warm or damp to touch, no sweat or wet spots visible Wet spots visible, no droplets or streams Excessive sweating, may include streams or droplets	0 1 2 3
Tail flicking (do not count flicking to chase off insects)	No tail flicking Occasional tail flicking (1-2 times/5 min) and/or holds tail away from body Frequent tail flicking (3-4 times/5 min), may hold tail away from body Excessive tail flicking (>5 times/5 min)	0 1 2 3
Kicking at abdomen (= with hind leg)	Quietly standing, no kicking Occasional kicking at abdomen (1-2 times/5 min) Frequent kicking at abdomen (3-4 times/5 min) Excessive kicking at abdomen (>5 times/5 min)	0 1 2 3
Pawing at floor (number of episodes)	Quietly standing, does not paw at floor Occasional pawing at floor (1-2 times/5 min) Frequent pawing at floor (3-4 times/5 min) Excessive pawing at floor (>5 times/5 min)	0 1 2 3
Head movements	No fast movements, head mostly at same height/in same direction Occasional head movements laterally or vertically, looking at flank (1-2 times/5 min) Frequent and fast head movements laterally or vertically, looking at flank (3-4 times/5 min) Excessive head movements, excessive looking at flank (>5 times/5 min), biting at flank (>1 time/5 min)	0 1 2 3
Pain sounds	No audible signs of pain Occasional teeth grinding or moaning (1-2 times/5 min) Frequent teeth grinding or moaning (3-4 times/5 min) Excessive teeth grinding or moaning (>5 times/5 min)	0 1 2 3
Overall appearance, reaction to observer(s)	Quiet but alert, approaches/turns to observer Alert, no reluctance to move, obvious reaction to sounds and/or movements Restless, constantly moving, exaggerated reaction to sounds and/or movements Stupor: the horse is not moving, head is lowered, reluctance to move	0 1 2 3
Reaction to palpation of painful area	No reaction to palpation Mild reaction to palpation Resistance to palpation Violent reaction to (attempt to) palpation	0 1 2 3
Total		.../42

Annex II- Facial Expression Pain Scale

Observation:

* Use this pain scale to determine:

- Pain at admitting *T=0*
- If given NSAIDs at university clinic; score 0.5 - 1 h later, *T=0,5/1*

OR - If pain management is unknown/ was administered before arrival at clinic; score after 2-3 h *T=2/3*

- The first and second morning after arrival *T=1st and 2nd morning*.

* Observe the patient for 60 seconds and be aware of sounds produced by the patient.

* Sum the scores to a total Facial Pain Expression score.

Patients label

Date:
Time:
Observer:
Time frame: T = 0 T = 0,5/1 T = 2/3

T = 1st morning T =2nd morning

Head	Normal head movement/ Interested in environment Less movement No Movement	0 1 2
Eyelids	Opened, sclera can be seen in case of eye/head movement More opened eyes or tightening of eyelids. An edge of the sclera can be seen for 50% of the time Obviously more opened eyes or obvious tightening of eyelids. Sclera can be seen more than 50% of the time	0 1 2
Focus	Focussed on environment Less focussed on environment Not focussed on environment	0 1 2
Nostrils	Relaxed A bit more opened Obviously more opened, nostril flaring and possibly audible breathing	0 1 2
Corners mouth/ Lips	Relaxed Lifted a bit Obvious lifted	0 1 2
Muscle tone head	No fasciculation's Mild fasciculation's Obvious fasciculation's	0 1 2
Flehming and/or Yawn	Not seen Seen	0 2
Teeth grinding and / or moaning	Not been heard Heard	0 2

Ears	Position: Orientation towards sound/ clear response with both ears or ear closest to source Delayed / reduced response to sounds Position: backwards / no response to sounds	0 1 2
Total		.../18

Annex III - VAS score veterinarians (Dutch)

Datum:

Naam:

Ervaring/werkzaamheden kliniek:

Voor ons onderzoek naar het meetbaar maken van pijn bij acute koliek paarden willen we graag een vergelijking maken tussen de pijnscores door middel van CPS en 'op het oog van de dierenarts' door middel van het VAS(visual analogue scale) score systeem. Met dit score systeem is het de bedoeling dat u een waarde geeft in pijn op een schaal van 0 tot 10. De '0' staat voor geen pijn en de '10' voor de ergste pijn die er is. We willen u vragen om bij elk van de filmpjes een streepje te zetten op een lijn van 10 cm.

Filmpje 1:

0 10

Filmpje 2:

0 10

Etc..

Hartelijk dank voor de medewerking!
Anne-Marie Sas & Marlijn van Gent.
Onderzoeksstudenten onder begeleiding van Thijs van Loon en Machteld van Dierendonck

Annex IV - Questionnaire (Original version)

Equine Personality Test(Ijichi et al., 2013)

The questionnaire asks about your observations of your horse's behaviour and its interaction with you, other people and other horses. **It is important that you answer the questions as truthfully as possible** (warts and all!).

Below are pairs of words joined by a line. Please put a cross on the line at the point that best describes this horse. For example a fairly spirited rather than a steady horse might be scored as:

Spirited	-----------X---	Steady	Score code
Easy-going	---	Intolerant	A - R
Argumentative	---	Well-mannered	A
Anxious	---	Confident	N - R
Obedient	---	Wayward	A - R
Sluggish	---	Forward-going	E
Willing	---	Stubborn	A - R
Placid	---	Active	E
Gentle	---	Rough	A - R
Adventurous	---	Habitual	E - R
Excitable	---	Laid-back	E - R
Nervous	---	Calm	N - R
Spirited	---	Steady	E - R
Relaxed	---	Tense	N
Quiet	---	Restless	N
Friendly	---	Standoffish	F4 - R

Please circle the number that indicates how often you have seen this horse behave in the following manner:

	Never		Some-times		Always	Score code
1. When it has the opportunity, how often does this horse initiate interaction with you?	1	2	3	4	5	F4
2. When it has the opportunity, how often does this horse initiate interaction with other people?	1	2	3	4	5	F4
3. When it has the opportunity, how often does this horse initiate interaction with other horses?	1	2	3	4	5	F5

Questionnaire (Dutch)

De vragenlijst gaat over uw waarnemingen aan het gedrag van uw paard en de interactie van het paard met u, andere mensen en andere paarden. Het is belangrijk dat u de vragen zo waarheidsgetrouw mogelijk invult.

Hieronder staan woordparen verbonden met een lijn. Zet u alstublieft een kruisje op de plek dat uw paard het best omschrijft. Een meer levendig dan gestadig paard wordt bijvoorbeeld als volgt gescoord:

Levendig |-------------------x--| Gestadig

Makkelijk in de omgang |---| Onverdraagzaam

Tegendraads |---| Goed gemanierd

Angstig |---| Zelfverzekerd

Gehoorzaam |---| Eigenzinnig

Traag |---| Voorwaarts

Gewillig |---| Koppig

Rustig |---| Actief

Zachtaardig |---| Ruig/onbehouwen

Avontuurlijk	|---|	Gewoon
Prikkelbaar	|---|	Beheerst
Nerveus	|---|	Kalm
Levendig	|---|	Gestadig
Ontspannen	|---|	Gespannen
Rustig	|---|	Onrustig
Vriendelijk	|---|	Afstandelijk

Omcirkelt u bij de volgende vragen alstublieft het nummer dat het gedrag van het paard met betrekking tot de volgende karaktereigenschappen het best beschrijft:

		Helemaal niet		gemiddeld		heel erg
1.	Hoe angstig is het paard over het algemeen in de buurt van andere paarden?	1	2	3	4	5
2.	Hoe energiek zou u dit paard over het algemeen noemen?	1	2	3	4	5
3.	Hoe betrouwbaar zou u dit paard over het algemeen noemen?	1	2	3	4	5

Annex V - Behavioral tests

Temperament testing - T2nd morning

Naam eigenaar:

Naam paard:

Test	How	Score categories	What is tested
Open Field /novel environment	VIDEO – First 2 minutes without observers in stable	OF	fearfulness
Sound			
Positive	Drop food pellets in a bucket (out of sight)	Sound	attentiveness
Negative	Loud noise (whistle) out of sight	Sound	fearfulness
Human approach			
Acute	VIDEO – First reaction to clinician, assistant or observer entering stable	HA	reactivity
Novel Object			
Colored item (wallet)	Stand 2 meters in front of horse and offer the object (see notes below). Watch reaction for 5 seconds.	NO	attentiveness
Touch sensitivity			
Skin sensitivity	Lightly trace a pen or finger along the horse's flank from shoulder to loins.	touch	sensitivity
Reactivity			
Garbage bag	The horse eats for 10 sec from a bucket on a garbage bag, then move the garbage bag (see notes below). Watch reaction	reaction	reactivity
Recovery			
Eating from bucket	Measure time until eating again after moving the garbage bag	reaction	attentiveness

OF: classification **Open Field**

1. Exploratory: horse calmly enters stall; smells bedding and walls; walks around
2. Neutral: horse stands still
3. Evasive: horse stays alongside the walls and does not stand in the middle of the stall

Sound: classification **Sound Tests**

1. Interested: horse raises head; looks in direction of sound; may move head in direction of sound
2. Neutral: does not appear to react, may move one ear
3. Fear: raises head; pulls/jumps away from noise; acute reaction

HA: classification **Human Approach** (*Welzijnsmonitor Paardenhouderij*)
1. Interested: horse moves toward person in a friendly manner; looks up or around; stretches neck toward person; smells person or touches them
2. Neutral: horse does not move in the direction of person; only turns ears toward person without turning the head
3. Threaten: horse does not clearly move toward or away from person, but does not appear friendly (flattens ears, bares teeth, threatens to kick etc.)
4. Aggressive: horse makes a clearly threatening and/or aggressive movement toward the person
5. Evasion: horse makes a clearly startled movement away from person

Notes: ***NO*** *offering:* ***Colored item- Novel Object*** *(Welzijnsmonitor Paardenhouderij)*
1. *Approach horse from the front in a calm manner: stop at 2 meters in front of horse with colored item in hand behind the back and click tongue twice.*
2. *Stretch arm with colored item forward, with the back of the hand facing upwards (item is not visible to the horse).*
3. *Turn the hand so that the colored item becomes visible.*
4. *Observe the (main) reaction of the horse for 5 seconds.*

NO: classification **Colored item (wallet)**
1. Touching: the horse immediately stretches its neck and possibly takes a step forward and touches the wallet within 5 seconds.
2. Reaching: the horse moves forward, carefully stretches the neck or may take a step forward; can smell the wallet but does not touch or does not touch within 5 seconds.
3. Neutral: the horse does not stretch its neck and does not take a step forward.
4. Turn away: the horse turns (head and/or entire body) calmly away from the wallet.
5. Evasion: the horse flinches, steps back/is scared of the wallet.

Touch sensitivity: classification **Touch** (adapted from *Welzijnsmonitor Paardenhouderij*)
1. No indication of tension and/or sensitivity (light tensing followed by relaxation of back muscles at touch is considered normal).
2. Indication of heightened sensitivity: slight dodging/ “hard” back / tensing; *no* or mild behavioral reaction (such as turning ears backward).
3. Indication of (very) high sensitivity: mild to severe reaction to pressure; pulling away from the touch, back flinching, kicking or an attempt to, biting, turning away from the researcher, walking away

Notes: ***Reactivity:*** *Eating from a bucket on a* ***garbage bag.*** *(adapted from Christensen et al 2008)*
1. *Let the horse eat from a bucket on a garbage bag for 10 sec*
2. *remove the garbage bag from underneath the bucket*
3. *Observe the reaction*
4. ***Recovery:*** *measure time until eating again from the bucket for 1 minute*
5. *If no reaction: wave the garbage bag near the ear of the horse*

Reactivity: classification **Reaction**
1. None -Horse does not react, chewing is not interrupted

2. Head up - Horse raises its head from the bucket, chewing may be briefly interrupted, not alert, not away
3. Alert - Horse stands vigilant with elevated neck, tail may be elevated, head and ears oriented towards test stimulus, chewing is interrupted, horse may move <u>up to 2 steps</u> away from the bucket
4. Away - Horse moves <u>3 or more steps</u> away from bucket in response to stimulus, followed by alertness
5. Flight - Horse <u>turns/jumps</u> away from bucket in a sudden movement, followed by trotting, alertness and possibly snorting

Recovery: classification **Reaction**

1. Doesn’t stop eating
2. Stops eating, starts eating again in <5 sec
3. Stops eating, starts eating again in 5-30 sec
4. Stops eating, starts eating again in 30-60 sec
5. Doesn’t start eating

Annex VI – form of consent

ONDERZOEK NAAR PIJN TEN GEVOLGE VAN KOLIEK

Geachte heer/mevrouw,

In het Departement Gezondheidszorg Paard wordt onderzoek gedaan naar pijn en optimale pijnbestrijding. Op dit moment zijn wij met de hulp van twee Master-studenten, Anne-Marie Sas en Marlijn van Gent, bezig om de hoeveelheid pijn die een paard met acute koliek ervaart objectief vast te stellen. Zo hopen we een objectieve 'pijnschaal' te kunnen ontwikkelen. In de toekomst kan deze manier van pijn bepalen er aan bijdragen dat de hoeveelheid pijnstillers beter kan worden afgestemd.

Voor dit onderzoek willen wij graag uw paard tijdens zijn of haar verblijf op onze kliniek observeren en op video vastleggen. Uw paard zal hiervan niets merken. Alle gegevens worden volledig anoniem verwerkt. Graag willen wij hiervoor uw toestemming vragen.
Ook willen wij u vragen, terwijl de dierenarts uw paard onderzoekt, om een paar vragen te beantwoorden (zie bijlage).

Met vriendelijke groeten,

Dr. Thijs van Loon
Specialist anesthesiologie
j.p.a.m.vanloon@uu.nl / 030 2531111

Dr. Machteld van Dierendonck
Gedragsdeskundige

Naam eigenaar: …………………………………………………………………….

geeft hierbij toestemming dat zijn/haar paard …………………………….
op video wordt vastgelegd voor onderzoek naar pijnbestrijding

Datum ……..………….. Handtekening ....……………………….