

https://dspace.org/

Hold your horses: the effect of play behaviour in horses (Equus caballus) under imposed stress

Howard, Jodi Anne

2024
University of the Witwatersrand, Johannesburg
https://hdl.handle.net/10539/42153
Howard, Jodi Anne. (2024). Hold your horses: the effect of play behaviour in horses (Equus caballus) under imposed stress [Master's dissertation, University of the Witwatersrand, Johannesburg]. WireDSpace.https://hdl.handle.net/10539/42153

Downloaded from a DSpace repository. https://dspace.org/



Hold your horses: the effect of play behaviour in horses (*Equus caballus*) under imposed stress

Jodi Anne Howard, student no. 2095526

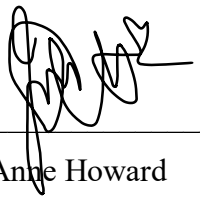
Master of Science thesis in the school of Animal Plant and Environmental Science

University of the Witwatersrand

February 2023 to March 2024

DECLARATION

I declare that this dissertation is my own unaided work. It is being submitted for the degree of Master of Science at the University of the Witwatersrand, Johannesburg. It has not been submitted before for any degree or examination at any other university.



Jodi Anne Howard

17 May 2024

Date

TABLE OF CONTENTS

ACKNOWLEDGEMENTS---------- 3
ABSTRACT---------- 4
KEYWORDS---------- 4
RATIONALE---------- 5
INTRODUCTION---------- 6
Stress---------- 6
Play---------- 8
Stress and Play----------10
Stress and Cognition----------10
Horses----------11
AIMS, OBJECTIVES, HYPOTHESES AND PREDICTIONS----------13
MATERIALS AND METHODS----------15
Study sites----------15
Study subjects----------15
Food choice test----------16
Habituation to experimental set-up----------16
Treatments----------17
Baseline treatment----------17
Stress treatment----------22
Play treatment----------22
Salivary cortisol measurements----------23
Data analysis----------24
RESULTS----------25
Behavioural diversity----------25
Frequency of behaviours----------30
Latency----------35
Emotional states----------40
Positive----------40
Negative----------41
Neutral----------44
Initial and final decision success----------44
Initial decision----------46
Final decision----------47
Salivary cortisol----------49
DISCUSSION----------51
General behaviours----------51
Decision-making----------54
Stable conditions----------55
Sex effect----------55
Salivary cortisol----------56
CONCLUSION----------57
REFERENCES----------59

ACKNOWLEDGEMENTS

First and foremost I would like to thank my supervisor Prof. N. Pillay, without whose guidance, encouragement and financial support, I would not have been able to complete my dissertation. It has been an incredible experience to work with him. Second, I would like to thank the owners of the Weltmeyer Stables, as well as the owners of the River Farm, for allowing me to use their horses and premises for my study. Thirdly, I would like to thank Lancet Laboratories for their assistance in obtaining the salivary cortisol levels from my samples. Fourth, I would like to thank my husband, Mr J.D. Howard, and my father, Dr D.A. Mol, for their assistance with putting the Y-maze together. Fifth I would like to thank my husband, Mr J.D. Howard, my mother, Dr T.M. Mol, and my friend, Ms Z.A. Viljoen for their assistance with my field work and for their continuous support and encouragement throughout my study. Sixth, I would like to thank my proofreaders for their assistance, I am immensely thankful for you all. Seventh, I would like to thank my extended family, friends and church community for their unending love and support throughout my Master's degree, I treasure you all very much. Last and most certainly not the least, I need to give all the glory to God for allowing me this opportunity and for revealing Himself through His creation throughout my Master's degree. I am because He is!

ABSTRACT

In animals, stress is inevitable, often resulting in behavioural and physiological responses to stressors. Due to the negative connotation of stress and positive connotation of play behaviour, stress and play are concepts not often coupled in research. A few studies have shown a negative association between play and stress whereas others have shown a positive association. Horses (*Equus caballus*), especially those in equestrian sports, are prone to acute and chronic stress. My study aimed to ascertain whether and how play behaviours can ameliorate acute stress in horses from two different stabling and working conditions (i.e. housed in a stable or on a farm). My study involved investigating the behaviours and decision-making (cognition) of the individual horses in a Y maze for a food reward. All horses were exposed to three treatments in sequence, following a repeated measures design: baseline – without a stressor; stress - exposing the horses to a short-term acoustic stressor; and play - exposing horses to a short term acoustic stressor and then initiating object play with them. Treatments were repeated twice. I recorded the behaviours of the horses and later extracted the following variables: behavioural diversity (i.e. number of behaviours displayed/total potential behaviours); frequency of behaviours (number of behaviours per 10 seconds); emotional states (positive, negative and neutral), and the time taken (latency) for the horse to make a decision in the maze. I recorded whether they made decisions at two points: 1) the decision point in the maze and 2) for a food reward located at one terminal arm of the maze; the two decisions could be uncoupled. I collected saliva from each horse before and after each treatment for later analysis of cortisol levels. I found that the horses in the stress treatment showed greater diversity and frequency of behaviours linked to a negative emotional state as well as poor decision-making compared to the play and baseline treatments. Overall, decision-making was better in the play treatment than in the baseline and stress treatments. The horses at the stable, housed under more constrained conditions were more stressed than the horses on the farm. Male horses also appeared to be more stressed compared to the female horses. Salivary cortisol was not a significant predictor of any behaviours and decision-making, nor was it predicted by treatments, stabling or sex. I found that play is able to ameliorate stress behaviourally and that it improves decision-making, which indicates that play behaviour could be a useful indicator of animal welfare. It is, however, also evident that the behavioural and physiological response to a stressor appears to be decoupled in the horses selected for study.

KEYWORDS

cognition, decision-making, horses, play, salivary cortisol, stress

RATIONALE

Horses have been domesticated for many centuries. As a result, horses have been trained in various equestrian sports, many of which require high-performance outputs from the horse. Such high-performance often leads to increased stress, and although stress can be beneficial in the short term (positive stress), it is not beneficial over long periods (negative stress). Due to the nature of high-performance equestrian sport, horses are under pressure to perform and therefore, under stress more often. Human intervention is thus necessary to ensure that the horses perform at their best regardless of levels of stress. As one can imagine, human intervention is not always in the best interest of the horse, and so, what if there was a way to reduce a horse's negative stress by natural means? What if play behaviour could reduce the negative stress that a horse experiences when it needs to perform? My study involved using play behaviour to decrease stress in horses, which could also have implications for other species experiencing some form of stress in a captive environment.

INTRODUCTION

Stress

Non-human animals, whether domesticated or free-living, undergo stress at some point in their lives. To understand what causes stress and the effects that stress can have on an animal, it is important to define stress. Stress consists of two components, the stressor and the stress response. A stressor, which can be intrinsic or extrinsic, can trigger a stress response (Selye 1936; Sapolsky 2002; Oliveira *et al.* 2010; Koolhaas *et al*. 2011; Bartolomé and Cockram 2016). Stressors can be physiological (e.g. hunger), behavioural (e.g. presence of unfamiliar humans) or physical (e.g. fatigue) (Sapolsky 2002). When an animal experiences a stressor, its brain determines whether the stimulus is perceived as a threat or not (Möstl and Palme 2002; Bartolomé and Cockram 2016; Esch *et al.* 2019). When a stressor is perceived as a threat, there will be various behavioural and physiological responses (Sapolsky 2002; Bartolomé and Cockram 2016).

A behavioural response, which can be species or sex-specific, involves fearful or aggressive behaviours in response to the stressor (Sapolsky 2002; Bartolomé and Cockram 2016). An example of species-specific behaviour was tested in three rodent species, Mongolian gerbils (*Meriones unguiculatus*), Wistar rats (*Rattus norvegicus*) and golden hamsters (*Mesocricetus auratus*) exposed to Brazilian constrictor snakes (*Epicrates cenchria assisi* and *Boa constrictor amarali*) in a complex labyrinth (Guimarães-Costa *et al*. 2007). The rats and hamsters showed more panic-like and fearful behaviours compared to the Mongolian gerbils which showed more exploratory behaviours along with the panic-like and fearful behaviours (Guimarães-Costa *et al*. 2007). This example shows the nature of the stress response where a prey species will generally react more fearfully in the presence of a predator or could react in a way to avoid the predator altogether as seen in the example of the Mongolian gerbils. As an example of sex-specific behaviours, male common water striders (*Gerris gracilicornis*) were bolder in the presence of predators compared to female water striders (Han, *et al.* 2015). In another study, male dog whelk (*Nucella lapillus*) expended more energy in the presence of predators compared to female dog whelks (Donelan and Trussel 2020). While these studies consider more specifically the personality of male and female animals, Mowrey and Portman (2012) showed that underlying neural mechanisms of biological sex, plays a significant role in behavioural decision-making where animals need to decide how to respond in various scenarios ranging from mate choice to predator avoidance, as seen in the roundworm (*Caenorhabditis*

elegans) and common fruit fly (*Drosophila melanogaster*). From this, it is evident that while personality plays a role in sex-specific decisions, the underlying neural mechanisms of biological sex also play a role in decision-making. A physiological response involves the autonomic nervous system which is more commonly referred to as the fight or flight response (Sapolsky 2002; Bartolomé and Cockram 2016). The sympathetic nervous system (SNS) is responsible for the release of catecholamines which increases the free glycogen and fatty acids in the blood as well as blood flow through muscles, which are the underlying mechanisms for particularly the flight response (Möstl and Palme 2002; Grace *et al*. 2009; Koolhaas *et al*. 2011; Bartolomé and Cockram 2016). A neuroendocrine response occurs via the hypothalamic pituitary adrenal (HPA) axis, causing the release of adrenocorticotrophic hormone (ACTH) which stimulates the release of cortisol from the adrenal cortex into the bloodstream (Möstl and Palme 2002; Grace *et al*. 2009; Koolhaas *et al*. 2011; Bartolomé and Cockram 2016; Kang and Lee 2016). Cortisol can affect metabolism (e.g. increased levels of cortisol can lead to weight gain; Christiansen, *et al.* 2007), the immune system (e.g. elevated cortisol levels can lead to increased stress hormones and therefore increased inflammatory cytokines which negatively affect the immune system response; Morey *et al.* 2015) as well as reproduction (e.g. increased levels of cortisol inhibit the gonadotropin-releasing hormone; Whirledge and Cidlowski 2010), and is often referred to as the "stress hormone" (Möstl and Palme 2002; Grace *et al*. 2009; Koolhaas *et al*. 2011; Bartolomé and Cockram 2016). Thus, when an animal is under stress, there is an increased level of cortisol and other stress hormones (glucocorticoids, catecholamines, cortisol) in its system, as seen in Arctic ground squirrels (*Urocitellus parryii*) that responded to a stressor (live trapping) with increased levels of plasma cortisol (Boonstra *et al*. 2001). In another example, red drum fish (*Sciaenops ocellatus*) displayed an increase in cortisol and glucose levels, when handled in shallow water (Thomas and Robertson 1991).

A stress response is often considered to be negative for human and non-human animals, but this is not always the case. In some instances, stress can be positive, for example, in racehorses. Before a race, horses that experience short-term or "positive" stress because of the release of stress hormones, can show enhanced performance (Bartolomé and Cockram 2016). Another example is in prey being hunted which are naturally more alert or experience a form of "positive" stress when there is a predator in the environment. With the increased secretion of hormones from the adrenal gland, the prey is more likely to behave or react in a way to ensure its survival when it encounters the predator (Apfelbach *et al*. 2005; McEwen and

Karatsoreos 2020). This is a short-term stress response and is necessary. For example, Sprague-Dawley rats (*Rattus norvegicus*) froze in the presence of a predator (Wallace and Rosen 2000), which is an immediate and short-term stress response. The problem, however, is when the stress response lasts for a long time, which then compromises the welfare of the animal. Long-term stress, also known as chronic stress, is accompanied by elevated levels of stress hormones, like cortisol, in the blood for longer periods, leading to behavioural and physiological changes (Bartolomé and Cockram 2016). O'Connor *et al*. (2010) showed that growing pigs (*Sus scrofa*) that were exposed to higher levels of atmospheric ammonia and were therefore under negative stress, had physiological changes, in which levels of hormones like cortisol fluctuated, which could impact positive behavioural experiences later in life.

Play

Play is a complex behavioural phenomenon (Burghardt 2005; Oliveira *et al.* 2010; Hausberger *et al.* 2012). Play is said to have a social, cognitive and even a developmental role in young children. Yet, the definition of play is still unclear, and there is uncertainty about how to define play (Bekoff 1984). Burghardt (2005) considered several aspects to define play based on Niko Tinbergen's (1963) four questions on how the mechanism, development, function and evolution can be considered to provide a different perspective of play behaviour. The problem, however, is that there are many aspects to consider within each of Tinbergen's questions, and one would need to address all of these questions before generating a widely accepted definition. Even though there was limited knowledge of play behaviour in non-human animals (hereafter animals), Burghardt, using the framework developed by Clark and Miller (1998) on child play, suggested five criteria for play in animals. These are that 1) play is not serious, 2) play is voluntary, 3) play is temporally or structurally different from other behaviours, 4) play is repeated, and 5) play is present when the animal is not under stress (Burghardt 2005). These criteria have been most helpful in past studies when identifying play behaviour in animals (Oliveira *et al.* 2010). Burgardt *et al.* (2015) used the five criteria to show that Cichlid fish (*Tropheus duboisi*) displayed forms of object play that were highly repetitive. Hausberger *et al.* (2012) used the five criteria to define play behaviour in riding school horses (*Equus caballus*) under stress and found that domesticated adult horses do exhibit play behaviours, but that this behaviour can be related to their long-term stress levels which is a form of negative stress.

Researchers have identified three forms or types of play, namely social, locomotor and object play (McDonnell and Poulin 2002; Burghardt 2005; Oliveira *et al.* 2010; Held and Špinka 2011). Social play is directed at another individual of the same species, and a good example is seen in play-fighting behaviour that often involves wrestling, nipping and rolling around (Burghardt 2005; Pellis and Pellis 2017). Play fighting occurs in a wide range of mammals. For example, juvenile rats (*Rattus norvegicus*) engage in play fighting by one rat approaching another rat and pouncing on it, the other rat will then roll over to get the attacker off of its neck, and so the behaviour continues (Vanderschuren *et al.* 1997; Pellis and Pellis 2017; Antonevich *et al.* 2019). Young felids chase, nip and wrestle during play (Antonevich *et al.* 2019). Locomotor play involves running or jumping and moving the head in various directions (also referred to as locomotor-rotational play) (Wilson and Kleinman 1974; Burghardt 2005). It can also be observed when the animal makes sudden changes in its direction of movement as it shakes or rotates its head (Wilson and Kleinman 1974; Burghardt 2005). This is most often observed in young animals like calves and foals (Wilson and Kleinman 1974; Burghardt 2005). For example, foals often run around and leap in the air, a behaviour coupled with bucking, where all four legs are off the ground at the same time as the foal jumps upward with its head down (McDonnell and Poulin 2002; Burghardt 2005; Held and Špinka 2011). Domestic pigs display locomotor play behaviour when they pivot or jump on the spot to face a different direction (Donaldson *et al*. 2002). Object play refers to play with an inanimate object (McDonnell and Poulin 2002; Burghardt 2005; Oliveira *et al.* 2010; Held and Špinka 2011). An example of object play was reported in horse foals by McDonnell and Poulin (2002), where the foals would pick up sticks or small branches and carry them around using their mouths. On occasion, it was also observed that they would throw the stick into the air and then pick it up again, repeatedly (McDonnell and Poulin 2002).

Play behaviour is most common in young animals. Thus, researchers have suggested that play behaviour has a developmental role, which explains why juveniles play more than adults (Waring 2003; Burghardt 2005). However, play is not completely absent in adult animals, and is most often observed in a social setting. For example, Yamanashi, *et al.* (2018) found that adult chimpanzees (*Pan troglodytes*) played with one another just before they were fed. In another example, adult horses ran with one another, kicked out or bucked, and also nipped at one another, behaviours which were interpreted as play (Hausberger *et al.* 2012).

Stress and Play

Hausberger, *et al.* (2012) found that social play increased levels of stress. In another study, play behaviour increased when rhesus monkeys (*Macaca mulatta*) displayed social stress, for example, when there was a shortage of food (Loy 1970; Held and Špinka 2011). These studies showed a positive link between play behaviour and stress and seemed to contradict Burghardt's (2005) fifth criterion (above) which states that play behaviour is present when the animal is not stressed. These studies provide clear examples of animals that play more when they are stressed, but the main aim of these studies was to establish whether play behaviour could be used as a reliable indicator of animal welfare and not to contradict Burghardt's ideas. However, other studies showed that play behaviour occurs in animals that are not stressed. For example, Lonsdorf *et al.* (2014) found that there were reduced stress hormone levels in the faecal samples of chimpanzees that showed higher levels of play. Thus, the chimpanzees were not stressed yet exhibited play behaviour. In contrast, Pellis and Pellis (2017) showed that play fighting can occur in animals when they are not in a stressed state. These studies show the dynamic interplay between stress and play, but what effect do stress, and play have on an animal's cognitive abilities?

Stress and Cognition

Cognition refers to how an animal integrates information about its surrounding environment and the processes involved in deciding how to act (or not) based on that information (McEwen and Sapolsky 1995; Mendl 1999; Shettleworth 2001; Sandi 2013; Soravia *et al*. 2021). The important elements in cognition are the perception and interpretation of stimuli, learning, memory and decision-making (Shettleworth 2001; Dias-Ferreira *et al.* 2009). Cognition in mammals occurs in two parts of the brain, through the hippocampus and prefrontal cortex that work to integrate information (Sigurdsson and Duvarci 2016). The hippocampus is arranged along a longitudinal axis in the temporal lobe of the brain and is involved in learning, memory and spatial navigation, (Sigurdsson and Duvarci 2016). The entorhinal cortex is the input source for the hippocampus (Sigurdsson and Duvarci 2016). Adrenal steroids like glucocorticoids, secreted by the adrenal glands, are also necessary to modulate the ability to store and retrieve information in the hippocampus (Shettleworth 2001; Brubaker and Udell 2016; Schubiger *et al.* 2020). The prefrontal cortex is essential for higher-order cognitive processes as well as emotional regulation and receives input from the mediodorsal thalamus (Sigurdsson and Duvarci 2016).

The hippocampus extends monosynaptic projections to the prefrontal cortex or there are bidirectional interactions between the regions via the nucleus reuniens (Sigurdsson and Duvarci 2016). The neuroendocrine pathways involved in stress appear to be synergistic with some of the physiological cognitive processes. This means that stress is associated with cognition mechanistically (Mika, *et al.* 2012). The Yerkes-Dodson law states that an animal's cognitive performance is best when the animal is in an optimal state of stress (Yerkes and Dodson 1908; Mendl 1999). This would be when the animal undergoes positive or short-term stress when the cortisol levels in the blood are not maintained at a high concentration for a prolonged period of time. When cortisol levels remain in the bloodstream at high concentrations, cognitive processes can be hindered, including permanent neuronal loss (McEwen and Sapolsky 1995).

Dias-Ferreira *et al.* (2009) showed that chronic stress can cause a reorganisation of hippocampal-prefrontal interactions in rats (exact species unknown), so that cognitive processes like decision-making, in particular, will be greatly affected. In rodents, chronic stress hinders the ability to navigate in a maze (Dias-Ferreira *et al.* 2009; Sigurdsson and Duvarci 2016). Bowman et al. (2003) showed that chronic stress in rats (exact species unknown) hindered learning and memory. It is because of the relationship between stress and cognition that many researchers have used cognition as a way of determining whether the animal is undergoing short-term (positive) or long-term (negative) stress (Mendl 1999).

Horses

The horse (*Equus caballus*) is a magnificent animal that has been domesticated for many centuries. Horses are herbivorous and are social, herd animals. Its behaviours have been well documented over many years (McDonnell and Poulin 2002; Dai *et al*. 2019; Lundberg *et al.* 2020; Torcivia and McDonnell 2021). It is known for being intelligent as seen in various studies where its innovativeness or cognitive ability has been tested (Krüger and Heinze 2008; Esch *et al.* 2019). In their study, Krüger and Heinze (2008) showed that a horse observing a trainer and demonstrator horse doing something specific, would copy what the other horse would do, provided the demonstrator horse was reliable. Another example of intelligence is "Clever Hans", a horse that drew much attention when he was able to perform mathematical calculations by tapping on his stable door (Samhita and Gross 2013). Although it was later discovered that the trainer of Clever Hans played a role in his abilities, it did not detract from

the fact that he was able to perform what he was trained to do (Samhita and Gross 2013: Lapuschkin *et al*. 2019).

The use of horses in the equestrian industry has grown over the years (Krüger and Heinze 2008). Here, horses are used for leisure (e.g. ridden for fun) or they are used in sports, competing in various events such as dressage, show-jumping, cross-country competitions or racing (Bartolomé and Cockram 2016; Esch *et al.* 2019). It is unfortunate, however, that in some instances horses have been placed under stress because of the roles that they have in the equestrian industry. Even the horses used for leisure, such as those at riding schools, undergo negative stress and thus experience compromised welfare (Hausberger *et al.* 2012). Riding school horses suffer vertebral disorders when they are ridden often and by many different riders with different skill levels (Lesimple, *et al.* 2010). Hausberger *et al.* (2012) found that riding school horses with these chronic vertebral disorders were more likely to exhibit play behaviour than horses that did not experience this form of chronic stress. The researchers explained this with the use of Carr's recuperation theory which states that play behaviours can act as an escape from the stress that an animal experiences.

AIM, OBJECTIVES, HYPOTHESES and PREDICTIONS

Aim

This study aimed to ascertain whether and how play behaviours can ameliorate acute stress in horses (*Equus caballus*) from two different stabling and working conditions. My study involved three parts 1) exposing horses to a short-term stressor, 2) initiating object play of the horses, and 3) measuring their decision-making (cognition). The experiment involved 3 treatments: 1) exposing a horse to a decision-making test in the form of a Y-maze, 2) exposing horses to a short-term stressor, and measuring their decision-making (cognition); and 3) exposing horses to a short-term stressor, initiating play thereafter and measuring their decision-making (cognition).

My study included five objectives:

- **Objective 1:** To record the behaviour of individual horses exposed to three treatments, namely a baseline, stress and play.
  - *Hypothesis:* The horses will change their behaviours before and after the stressor and during play. I expected that the horses would be in their normal (relaxed) state prior to exposure to the stressor (baseline) but would show signs of fearfulness and anxiety after the stressor was introduced (stress treatment) and show reduced stress when a play object was introduced (play treatment).

- **Objective 2:** To establish, using a cognitive test, whether play behaviour reduces the effects of a short term stressor and influences decision-making.
  - *Hypothesis:* I expected that the decision-making of the horses in locating a preferred food would be better following a play treatment than a stress and no play treatment. Decision-making in the play treatment would be similar to that of the baseline treatment.

- **Objective 3:** To investigate how two groups of horses, with different stabling conditions (horses stabled and attended to daily versus horses not stabled and not attended to often) and histories, respond to the treatments.
  - *Hypothesis:* The horses stabled under more constrained conditions will display higher stress as compared to the horses under less constrained conditions. For this reason, the

horses under less constrained conditions will be more likely to make the correct decision in the maze compared to the horses under more constrained conditions.

- **Objective 4:** To investigate whether sex influences the behaviour expressed by the horses during the treatments.
  - *Hypothesis:* Female horses will display more stress-like behaviours than male horses. For this reason the male horses will be more likely to make the correct decision in the maze compared to female horses.

- **Objective 5:** To assess the change in the level of salivary cortisol of the horses prior to the stressor (baseline), after the stressor (stress treatment) and after play (play treatment).
  - *Prediction:* The level of cortisol in the saliva will increase when the horse is stressed and will be lower after the play versus when no-play behaviour takes place.

MATERIALS AND METHODS

Study sites

Horses at two sites were selected for study, both located in the Gauteng Province, South Africa, approximately 54 km apart. The first site was the Weltmeyer stables (26°16'00"S 28°04'56"E) located in the south of Johannesburg just off the N12 highway. This study site was surrounded by houses and was a horse farm open to the public. The second site was the River or Heidelberg farm (26°38'22"S 28°13'20"E) based close to the Suikerbosrand Nature Reserve. There were very few houses surrounding this farm and the farm was privately owned. I obtained permission to conduct my study at these two sites. My experiments were approved by the University of the Witwatersrand. The Animal Research and Ethics Committee (AREC) clearance numbers are AREC23-03-018B and AREC23-06-007B.

Study subjects

Twenty horses (*Equus caballus*) were used in this study, 10 from the Heidelberg farm and 10 from the Weltmeyer stables. Three horses died during this study, two from the Heidelberg farm due to African Horse sickness and one from the Weltmeyer stables due to a fractured leg that turned septic. The horses at the Weltmeyer stables were 7-24 years of age. These horses comprised of five thoroughbreds (one mare and four geldings), three crossbreeds (two mares and one gelding) and one Clydesdale (gelding). They were all riding school horses and bred on various farms (exact information unknown). They were fed three times a day (foods of various nutritional value like lucerne chaff, maize, barley, oats and molasses), spent the mornings grazing in the paddock and were ridden in the afternoons by various riders (each horse was ridden for a maximum of 1.5 hours in the afternoons, including show-jumping, dressage techniques and flat work) and stabled individually in the evening. These horses were regularly groomed, shoed and checked by the veterinarian. The horses at the Heidelberg farm were 5-16 years of age. These horses comprised of Boer ponies (two mares and six geldings) and were bred on various farms (exact information unknown). They were not ridden often and were considered working horses (i.e. they were used to patrol fence lines and herd cattle). These horses were not fed during the day and so their main food source was pasture grass. They were, however, fed a maize supplement once every few days during winter. When the horses were not ridden, they grazed in the field. They were not stabled at night and were not often groomed or checked by the veterinarian. None of these horses were shoed.

Food choice test

Prior to experiments, each horse underwent a food choice test individually. The horses were presented with one cup (250ml) each of three food types, carrots, apples and horse lucerne pellets (selected because they were considered treats for horses; personal observation), each in their own separate plastic feeding bowl (25cm by 15cm by 10cm). The feeding bowls, of different colours randomly assigned for each food test, were placed one metre apart, and each horse was given 10 minutes to sample the three food types. This duration was chosen based on my experience with horses over many years. The food type that was consumed and finished first was regarded as the most preferred food type. The most preferred food type for each horse was then used as the incentive during the decision-making tests later on.

Habituation to the experimental set-up

Before any of the experiments, the horses were exposed individually to a Y-maze set-up (Clark and Smith, 2013; Rama *et al.*, 2016) that was used later for decision-making (see below). The Y-maze consisted of eleven wooden poles (1.5m in height) and rope (2 of 6m in length, 2 of 3.7m in length and 2 of 3m in length) (Figure 1). The wooden poles were sunk into the ground and the ropes were secured to the tops of the wooden poles with metal rings. The horse was individually led into the maze from the bottom of the "Y" and was then left to move forward by itself. This was done once a day for 15 minutes for two weeks before the start of the experiments to habituate the horses to the maze (Christensen *et al.* 2011; Leiner and Fendt, 2011).

The horses were also individually exposed to a 1m diameter ball used for play. This involved two weeks of training where the horse was placed in a round pen (16m in diameter) with the ball. Each horse interacted with the ball twice per week for 15 minutes (i.e. 4 time periods in total). The horse was encouraged to nudge the ball with its muzzle and was rewarded with its preferred food (see above), each time its muzzle touched the ball; this is a widely accepted training method for a horse to engage with an object (Tao of Horsemanship, 2020).

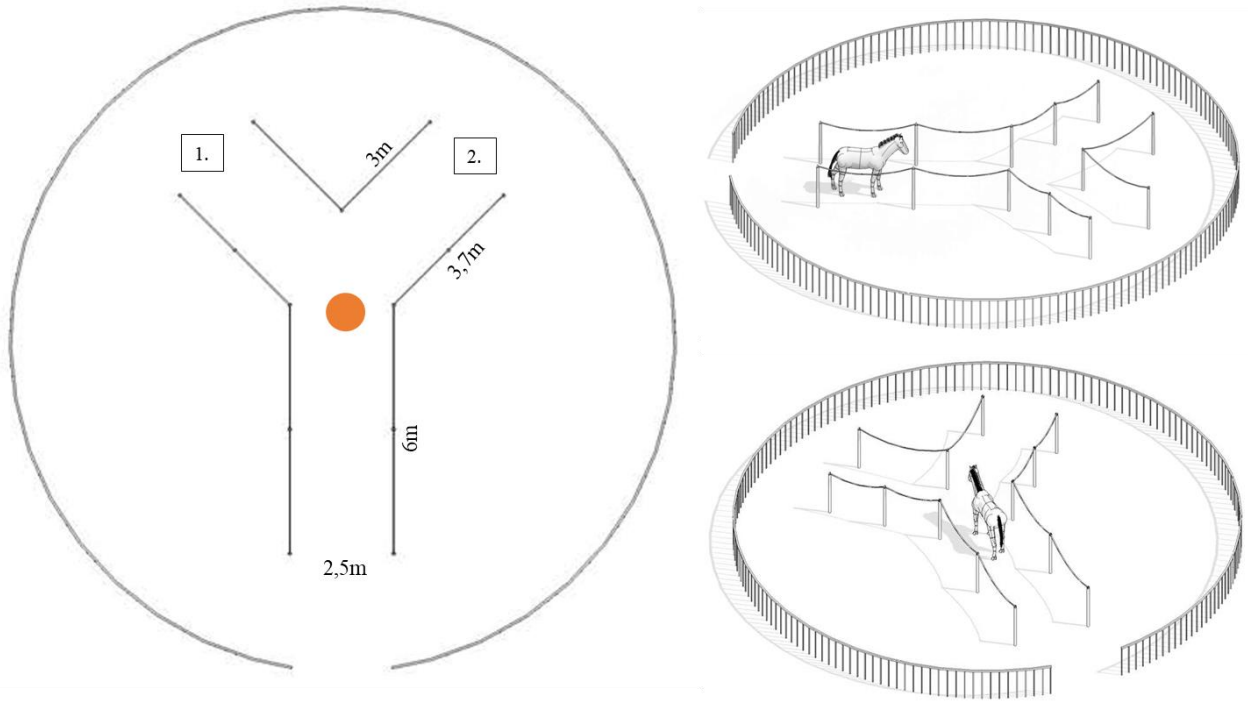


Figure 1. Schematic diagrams of the Y-maze within a round pen were used for the cognitive tests (produced using Autodesk Revit, 2018). The horse entered the maze at the bottom and moved toward the orange dot. The orange dot represents the decision point where the horse had to move to exit 1 or 2.

Treatments

Each horse underwent three treatments, namely baseline, stress and play. The horse's behaviours were recorded using a smartphone (Samsung Galaxy A50) in each treatment. On each of the experimental days, the horses first underwent a baseline treatment before undergoing the stress or play treatments. The stress and play treatments were randomly assigned to each horse with a minimum of two days between each treatment. Each horse was tested twice in the treatments, with a minimum of one week between repeat experimental treatments (Figure 2).

Baseline treatment

Each horse was placed in a round pen (16m in diameter), in which the Y-maze was already set up. The salivary cortisol of the horse was taken (see below). Thereafter, it was exposed to the Y-maze (already set-up), where one arm (either 1 or 2; Figure 1) had the horse's preferred food type and the other arm (either 1 or 2; Figure 1) had a feeding bowl with rocks

(the position of the food vs rocks randomly decided). The rocks were used in the feeding bowl of one of the arms as a neutral scent and not an incentive.

I recorded the time taken (latency in seconds) for the horse to reach the decision point before the split in the arms (Figure 1), the time to make a decision (head oriented into a particular arm), the first arm entered, and the time taken to reach the food in an arm (Table 1). Latency was recorded using a stopwatch. An initial decision was noted when the horse turned its head to face into one of the arms. A decision-making test was noted as successful when the horse chose the arm that contained its preferred food reward as the final decision. The behaviours (i.e. the position of its ears and tail, the posture of its head, neck and body and the movement of its legs; Table 2) were recorded continuously 5 minutes before (at the start of experiments in the round maze), during (in the maze), and 5 minutes after (in the round pen after exiting the maze), the experiment. From these behavioural recordings, the diversity of behaviours (i.e. the number of different behaviours shown during the experiment), frequency of behaviours (i.e. the number of times each behaviour was performed per unit time) and the emotional states of each of the horses were calculated (Table 1). Classification of the horse's emotional state was based on whether it showed 80% of the behaviours classified as positive, negative or neutral (Table 2). This means that if a horse showed 80% or more of the behaviours under the positive emotional state category then the horse would be classified as being in a positive emotional state. Similarly, if the horse showed 80% or more of the behaviours under the negative emotional state category then the horse would be classified as being in a negative emotional state. The same can be applied to a horse classified as being in a neutral emotional state.

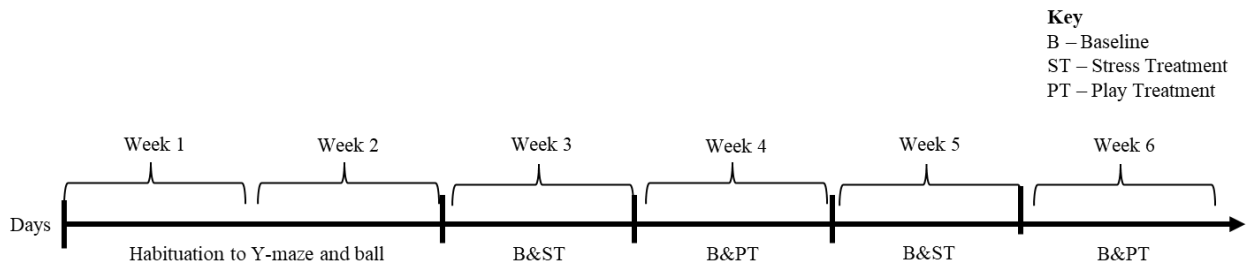


Figure 2. Timeline of events showing the habituation period of the horses to the Y-maze and the ball and when each experiment took place.

General behavioural observations of the horses in their paddocks were conducted before the start of the experiments for two hours on each farm. This was done to construct an ethogram for behaviour (Table 2).

Table 1: The dependent variables recorded in these and the method of calculation.

Dependent variables	How calculated
Behavioural diversity	The number of unique behaviours per 10 seconds out of a total of 26 behaviours, listed in Table 2, that the horse displayed when in the maze. From this, the behavioural diversity was grouped into diversity of behaviours before the decision point in the maze and diversity of behaviours from the decision point to the end point of the maze so that initial and final decision success were more easily distinguished.
Frequency of behaviours	The number of times each behaviour, listed in Table 2, occurred per 10 seconds. From this, the frequency of behaviours was grouped into frequency of behaviours before the decision point in the maze and frequency of behaviours from the decision point to the end point of the maze so that initial and final decision success were more easily distinguished.
Latency (seconds)	Total time taken from the point where the horse entered the maze to the point where the horse exited the maze. From this, the total latency was divided into latency from the start of the maze to the decision point and the latency from the decision point to the end point of the maze so that initial and final decision success were more easily distinguished.
Emotional states	Positive, negative and neutral states as seen in Table 2. A horse was said to be in a particular state when 80% of the behaviours associated with said state were observed (i.e. 80% of the behaviours in a column in Table 2).
Initial decision	An initial decision was recorded as successful when the arm of the maze with the preferred food type was selected or unsuccessful when the arm of the maze with the rocks in the feeding bowl was selected, depending on the position of the head.
Final decision	A final decision was recorded as successful when the arm of the maze with the preferred food type was selected or unsuccessful when the arm of the maze with the rocks in the feeding bowl was selected.

Table 2. Description of how the body posture of the horses was used to classify them into three emotional states. States utilised in this study were modified from ethograms from Lundberg *et al.,* (2020) and Dai *et al*., (2019).

Behaviour	Emotional states		
	Positive	Negative	Neutral
Position of ears	-Pointing acutely upward/outward -Often coupled with vocalizations (neigh)	-Flat back/ pointing acutely upward -Often coupled with aggressive or skittish behaviours	-Neither far or flat back nor pointing acutely upward, but rather slumped outward -Often coupled with grazing
Position of head and neck	-Raised/ arced toward ground/ rapidly swung from side to side playfully -Often coupled with bucking and upbeat movements	-Raised/ arced downward -Often coupled with jumping from side to side, kicking out	-Not lifted, nor arced towards the ground, Neck slightly stretched out
Position of tail	-Slightly lifted/ falls over hindquarters -Often coupled with upbeat movements	-Raised/ pulled between the hindquarters/ flicked rapidly from side to side -Often coupled with a frozen stance	-Neither lifted nor pulled between the horse's legs/ hanging down, swinging from side to side (in presence of flies)
Position of nostrils	-Not flared, do not appear big. Regular sniffing (often associated with playful behaviour and interaction with an object).	-Flared, nostrils appear big -Often coupled with rapid breathing	-Not flared, moves rhythmically as the horse breathes in and out
Position of eyes	-Pupils not dilated, do not appear wide -Often coupled with attentive behaviour	-Wide, dilated pupils -Often coupled with skittish movements. Horse is also very aware of surroundings (anxious behaviours)	-Pupils are not dilated, sometimes half closed as the horse rests
Position of body	-Aligned with the head and neck -Often coupled with bucking and upbeat movements	-Not aligned with the head and neck -Often coupled with jumping from side to side, kicking out	-Aligned with head and neck
Motion	-Upbeat, yet relaxed, walking (four-beat gait), trotting (two-beat gait, legs	-Upbeat, walking (four-beat gait, legs lifting high off the ground), trotting (two-beat gait,	-Relaxed, one back leg slightly bent as the tip of the hoof is the only part touching the

	diagonally synchronised), cantering (asymmetrical three-beat gait) -Often coupled with bucking	legs diagonally synchronised, legs lifting high off the ground), cantering (asymmetrical three-beat gait) -Often coupled with throwing head and neck forward, sudden halting (horse freezes in place), rapid breathing, jumping from side to side, kicking out	ground, walking (four-beat gait), trotting (two-beat gait, legs diagonally synchronised), cantering (asymmetrical three-beat gait) -Often coupled with grazing
Bucking	-All four legs lifted from off the ground simultaneously -Often coupled with arced neck or head and neck rapidly swung from side to side	-Horse does not buck when in a negative state, but may jump up and down or from side to side	-Horse does not buck in a neutral state
Rearing	-Stands up onto only hind legs -Often coupled with other playful behaviours	-Stands up onto only hind legs -Often coupled with ears flat against neck and kicking out, rapid breathing	-Horse does not rear in a neutral state
Kicking	-Horse does not normally kick out in a positive state	-Kicks out with one or two legs -Often coupled with rapid breathing	-Horse does not kick out in a neutral state
Pawing	-Horse does not paw at the ground when in a positive state	-Front leg lifted then rapidly extended in a forward direction as it hits the ground and is then dragged back against the ground	-Horse does not paw at the ground when in a neutral state
Grazing	-Muzzle touches the grass, chewing -Often coupled with more upbeat and frequent movements	-Horse does not normally graze when in a negative state	-Muzzle touches the grass, eating -Often coupled with slumped ears and a relaxed stance

Stress treatment (Figure 3)

After the baseline treatment in the round pen, the horse's behaviours were recorded continuously for 5 minutes (Table 2). A recorded motorbike and then a drone sound was played from a smartphone connected to a Bluetooth device (Volkano), for 5 and 5 seconds consecutively (Figure 3). The sounds were played for 10 seconds in total because from my experience with horses, the horses become stressed within 5 to 10 seconds of hearing a strange noise. The device was placed at random and out of sight near the round pen and the sounds were played at a maximum frequency of 20 kHz. The motorbike and drone sounds were selected because, in my observations during horse riding activities, the horses became anxious when a motorbike drove past their paddocks or when a drone was flown overhead. Immediately after the sounds, the horse underwent the decision-making test, as mentioned above. The horse's behaviours were recorded individually during the experiment and then continuously after the experiment for 5 minutes (Table 2). After 10 minutes (from when the horse was exposed to the stressor) the saliva of the horse was collected (see below). The latency, behavioural diversity, behavioural frequency, emotional states, initial decision and final decision were recorded (Table 1).

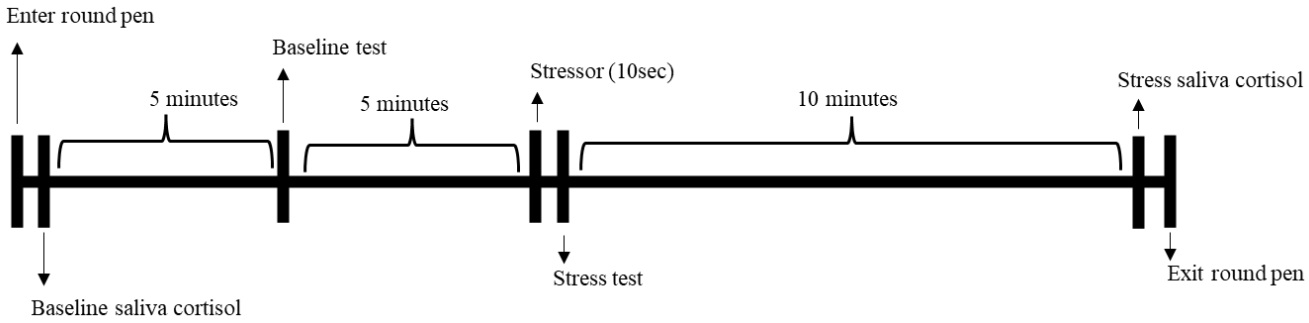


Figure 3. Timeline of the order of events on the day that the stress treatment took place. Each horse went through the stress treatment twice.

Play treatment (Figure 4)

After the baseline treatment in the round pen, the horse's behaviours were recorded continuously for 5 minutes (Table 2). A recorded motorbike and drone sound was then played from a smartphone for 5 and 5 seconds consecutively (Figure 4). The ball was then placed in the round pen. The horse was allowed to play with the object for 5 minutes, and its behaviour was recorded (Table 2). Thereafter, the horse underwent the decision-making test, as mentioned

above. The horse's behaviours were recorded individually during the experiment and then continuously after the experiment for 5 minutes (Table 2). After 10 minutes, from when the horse was exposed to the stressor, saliva was collected from the horse (see below). The latency, behavioural diversity, behavioural frequency, emotional states, initial decision and final decision were recorded (Table 1).

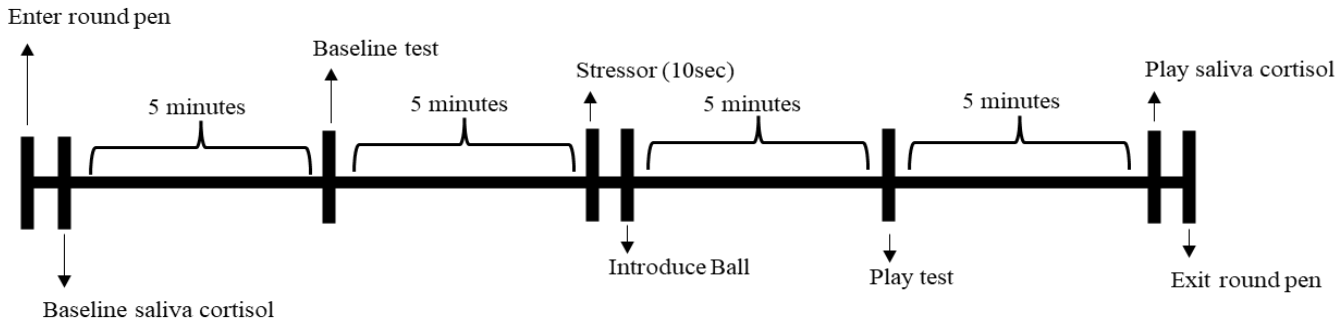


Figure 4. Timeline of order of events on the day that the play treatment took place. Each horse went through the play treatment twice.

Salivary cortisol measurements

Saliva sampling is one of the least invasive and simple ways of determining the levels of cortisol in animals (Sauer *et al.,* 2020). Saliva samples were collected from each horse before the start of the experiment and during the treatments. This was done to ascertain the level of cortisol present before and after the imposed stress in the horses. During the stress treatment, a saliva sample was collected while the horse was in the round pen and again 10 minutes after the horse had been exposed to the stressor. The time in between saliva sampling was chosen based on a study by Olczak *et al.,* (2018) who showed that increased levels of cortisol are present in the saliva of an animal within 10 minutes of exposure to a stressor. During the play treatment, a saliva sample was collected while the horse was in the round pen and again 10 minutes later after the horse had been exposed to the stressor and the ball.

I collected the saliva sample using a cotton swab under initial veterinary supervision. Using latex gloves, I placed the cotton swab into the horse's mouth and rubbed it against the horse's tongue and cheek for 30 seconds (Kang and Lee 2016). The sample was placed in a Salivette^{TM} tube, which is a cylindrical tube used to collect saliva for analysis (Kang and Lee 2016; Olczak *et al.* 2018). The tubes were immediately cooled at -20°C in a mobile fridge (Kang and Lee 2016; Olczak *et al.* 2018) and remained in the fridge until I delivered the

samples personally to Lancet Laboratories (Richmond, Johannesburg) for analysis, where they used the LC-MS/MS method to test for salivary cortisol (Sauer *et al.* 2020).

Data analysis

The video-recordings of the tests were scored manually onto data sheets. Scoring was done by me, and a subset was scored by another person who was naive to the experimental treatments (interobserver reliability – Personality: $0.98 < rs < 0.99$). All data analyses were done in R (R version 4.3.1 (2023-06-16 ucrt)). The data from both sets of repeats (treatments run twice) were included in the analyses. To test for normality, Shapiro-Wilk tests were run. None of the data was normally distributed ($p < 0.050$). Therefore, models were constructed using the glmm function in the DHARMa, lme4 and glmm packages. and statistical tests were performed using the ANOVA function in the car package. In these models, treatment, stable, sex, treatment x stable, and treatment x sex were fixed factors and salivary cortisol was the continuous predictor (covariate). Horse ID was included as a random factor. Separate models were run for the dependent variables, including behavioural diversity before and after the decision point, behavioural frequency before and after the decision point, latency before and after the decision point, emotional state, initial decision, and final decision. A separate model was also run for the salivary cortisol where the fixed factors were the same as above. Different families and link functions in the generalized linear mixed model were used for different dependent variables: family=Gamma, link=log for behavioural diversity, behavioural frequency, latency and salivary cortisol; family=Poisson, link=log for emotional state; and family=Binomial, link=logit for initial decision and final decision. Significant fixed effects were then tested using estimated marginal means (emmeans, magrittr, lme4 packages) for pairwise relationships.

Spearman Rank regressions were used to analyse the changes in behaviours before horses reached the decision point in the maze and after the decision point for these input variables: behavioural diversity, behavioural frequency, and latency. Regressions were not run for emotional states because the neutral emotional state was not autocorrelated and pulled the data in the opposite direction.

I set up contingency tables to produce matrices for the initial and the final decision success for treatment, stable and sex. Fisher's exact tests were run to determine if there was an association between the initial decision and the final decision for treatment, stable and sex (categorical variables).

RESULTS

The horses in this study, with different stabling conditions and histories, completed the experiments (baseline treatment, stress treatment, play treatment) twice. Their behaviours were recorded before and after they entered the Y-maze (already set-up in the round pen). The following parameters were measured during the experiments; behavioural diversity (before and after the decision point in the maze), frequency of behaviours per unit time (before and after the decision point in the maze), latency in seconds (before the decision point and after the decision point to when the horse exited the maze), emotional states (positive, negative and neutral), initial and final decision success of a test (Table 1). The results for each of the parameters follow below.

Behavioural diversity

Behavioural diversity was measured as the number of unique behaviours that the horse displayed. From this, the behavioural diversity was grouped into diversity of behaviours from the start to the decision point in the maze (diversity before) and diversity of behaviours from the decision point to the end point of the maze (diversity after).

Treatment was a significant predictor of behavioural diversity before the decision point (GLMM: $\chi^2 = 7.36$, df = 2, p = 0.025). Post hoc tests showed that the behavioural diversity in the stress treatment (median = 11.00) was greater than the baseline treatment (median = 10.00). There was no significant difference between the baseline and play treatments and the play (median = 11.00) and stress treatments for behavioural diversity before the decision point (Figure 5a). Stable had a significant effect on the behavioural diversity before the decision point (GLMM: $\chi^2 = 18.80$, df = 1, $p < 0.001$). Post hoc tests showed that the behavioural diversity before the decision point of horses at the Weltmeyer stables (median = 11.00) was greater than those on the Heidelberg farm (median = 10.00) (Figure 5b). There was a significant effect of sex on behavioural diversity before the decision point (GLMM: $\chi^2 = 9.44$, df = 1, p = 0.002). Post hoc tests showed that the behavioural diversity among the male horses (median = 11.00) was greater than that of female horses (median = 11.00) (Figure 5c). The interaction between treatment and stable was a significant predictor of behavioural diversity before the decision point (GLMM: $\chi^2 = 9.52$, df = 2, p = 0.009). Post hoc tests showed that the behavioural diversity of the horses at the Weltmeyer stables was greater than that of the horses at the Heidelberg farm in the stress treatment (median = 11.00 vs 10.00). There were no significant

differences in the baseline (median = 11.00 vs 10.00) and play treatments (median = 11.00 vs 10.00) (Figure 5d). There was no significant effect between treatment and sex on behavioural diversity before the decision point (GLMM: $\chi^2 = 0.53$, df = 2, p = 0.765). The median scores were as follows: baseline (females = 10.00 vs males = 11.00); stress (females = 11.00 vs males = 11.00); and play (females = 11.00 vs males = 11.00) treatments. Salivary cortisol was not a significant predictor of behavioural diversity before the decision point (GLMM: $\chi^2 = 0.78$, df = 1, p = 0.376).

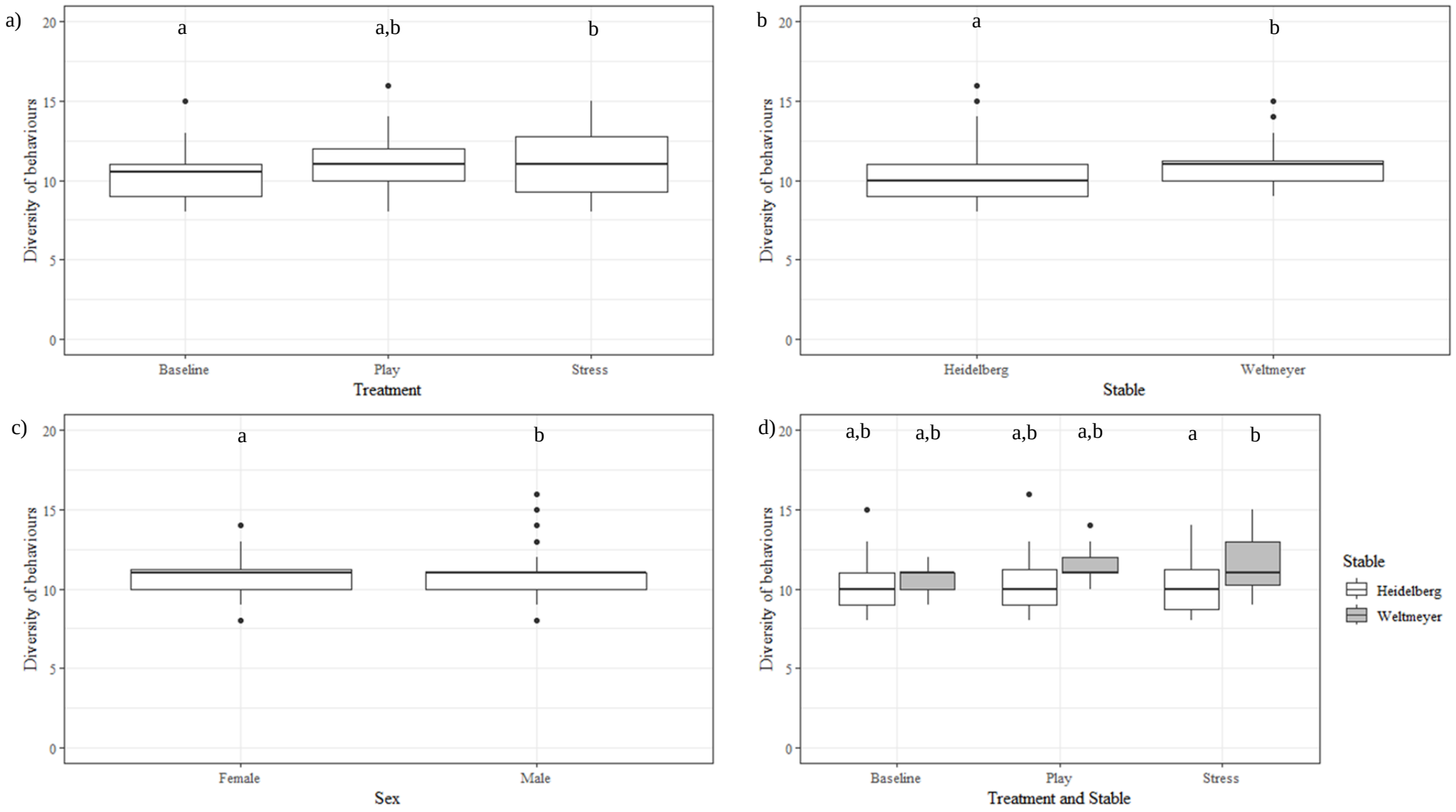


Figure 5. Behavioural diversity before the decision point a) in three treatments, b) at both stables, c) between female and male horses, d) interaction between treatment and stable. Boxplots indicate the median (central dark line), the 1st and 3rd interquartiles (boxes), minimum and maximum scores (whiskers) and outliers (dots). Bars with the same alphabets are not significantly different.

Stable had a significant effect on the behavioural diversity after the decision point (GLMM: $\chi^2 = 10.02$, df = 1, $p < 0.001$). Post hoc tests showed that the behavioural diversity after the decision point of the horses at the Weltmeyer stables (median = 11.00) was greater than those on the Heidelberg farm (median = 10.00) (Figure 6a). The interaction between treatment and stable was a significant predictor of behavioural diversity after the decision point (GLMM: $\chi^2 = 13.05$, df = 2, $p = 0.001$). Post hoc tests showed that the behavioural diversity, after the decision point, of the horses at the Weltmeyer stables was greater than those of the horses at the Heidelberg farm in the stress (medians = 11.00 and 10.00) and play treatments (medians = 10.00 and 9.50) (Figure 6b). There was no significant difference in the baseline treatment (medians = 11.00 and 10.00). Treatment was not a significant predictor of behavioural diversity after the decision point (GLMM: $\chi^2 = 1.81$, df = 2, $p = 0.405$). The median scores were as follows: baseline (11.00); stress (11.00); and play (10.00) treatments. There was no significant effect of sex on the behavioural diversity after the decision point (GLMM: $\chi^2 = 1.10$, df = 2, $p = 0.295$). Median scores were as follows: females (10.00); and males (11.00). There was no significant interaction between the horses at the Weltmeyer stables and the Heidelberg farm for the baseline treatment (medians = 11.00 and 10.00). The interaction between treatment and sex was not a significant predictor of behavioural diversity after the decision point (GLMM: $\chi^2 = 4.16$, df = 2, $p = 0.125$). The median scores were as follows: baseline (females = 11.00 vs males = 11.00); stress (females = 11.00 vs males = 11.00); and play (females = 10.00 vs males = 10.00) treatments. Salivary cortisol was not a significant predictor of behavioural diversity after the decision point (GLMM: $\chi^2 = 0.002$, df = 1, $p = 0.966$).

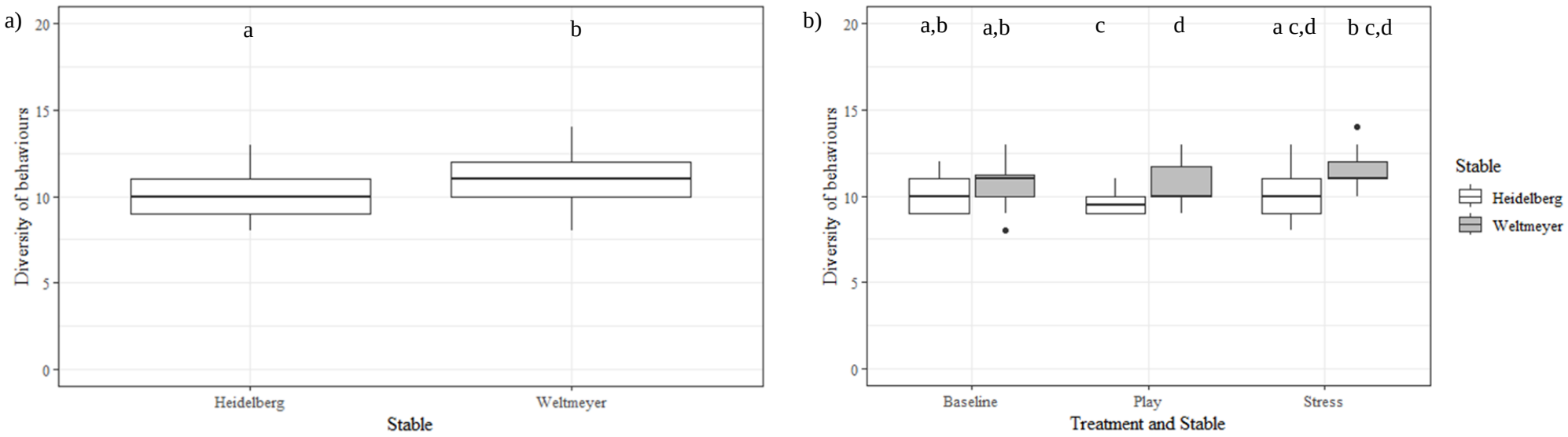


Figure 6. Behavioural diversity after the decision point a) at both stables, and b) on both farms in the three treatments. Boxplots indicate the median (central dark line), the 1st and 3rd interquartiles (boxes), minimum and maximum scores (whiskers) and outliers (dots). Bars with the same alphabets are not significantly different.

A Spearman Rank regression showed a significant positive relationship ($R = 0.45$, $p < 0.001$) between behavioural diversity before and after the decision point, indicating that the behavioural diversity of the horses was consistent when they entered and exited the maze (Figure 7).

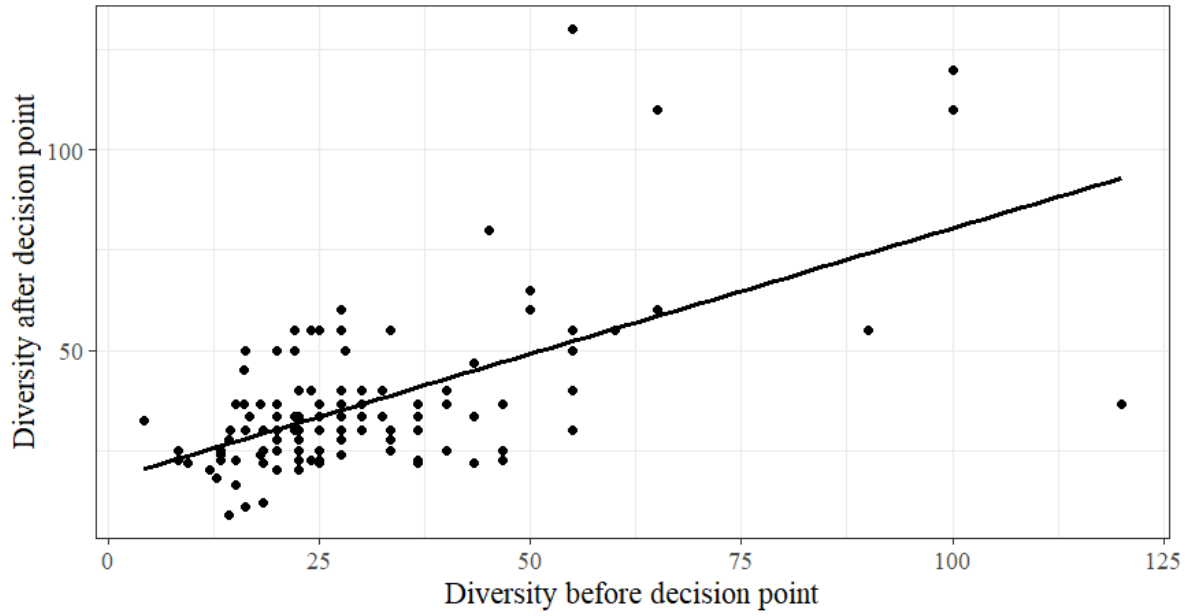


Figure 7. Regression of behavioural diversity before and after the decision point. The regression is significant ($p < 0.001$)

Frequency of behaviours

The frequency of behaviours was measured as the number of times a behaviour occurred per 10 seconds. From this, the frequency of behaviours was grouped into frequency of behaviours from the start to the decision point in the maze (frequency before) and frequency of behaviours from the decision point to the end point of the maze (frequency after).

Treatment was a significant predictor of the frequency of behaviours before the decision point (GLMM: $\chi^2 = 9.97$, $df = 2$, $p = 0.007$). Post hoc tests showed that the frequency of behaviours in the stress treatment (median = 32.50) was greater than the baseline treatment (median = 27.50). There was no significant difference between the baseline and play treatments and the play (median = 28.75) and stress treatments for the frequency of behaviours before the decision point (Figure 8a). A significant effect was observed of stable on the frequency of behaviours before the decision point (GLMM: $\chi^2 = 10.79$, $df = 1$, $p = 0.001$). Post hoc tests

showed that the frequency of behaviours before the decision point of horses at the Weltmeyer stables (median = 32.50) was greater than those on the Heidelberg farm (median = 24.00) (Figure 8b). The interaction between treatment and stable was a significant predictor of the frequency of behaviours before the decision point (GLMM: $\chi^2 = 9.46$, df = 2, p = 0.009). Post hoc tests showed that the frequency of behaviours of the horses at the Weltmeyer stables was greater than those of the horses at the Heidelberg farm for the stress treatment (median = 48.33 vs 20.00). There were no significant differences in the baseline (median = 30.00 vs 26.75) and play treatments (median = 31.00 vs 21.25) (Figure 8c). There was no significant effect of sex on the frequency of behaviours before the decision point (GLMM: $\chi^2 = 2.04$, df = 1, p = 0.154). The median scores were as follows: females (26.00) and males (30.00). There was no significant interaction between the horses at the Weltmeyer stables and those on the Heidelberg farm for the baseline treatment (median = 30.00 vs 26.75). The interaction between treatment and sex was not a significant predictor of the frequency of behaviours before the decision point (GLMM: $\chi^2 = 0.34$, df = 2, p = 0.842). The median scores were as follows: baseline (females = 25.50 vs males = 28.75); stress (females = 31.25 vs males = 36.25); and play (females = 25.00 vs males = 30.00) treatments. Salivary cortisol was not a significant predictor of the frequency of behaviours before the decision point (GLMM: $\chi^2 = 2.56$, df = 2, p = 0.110).

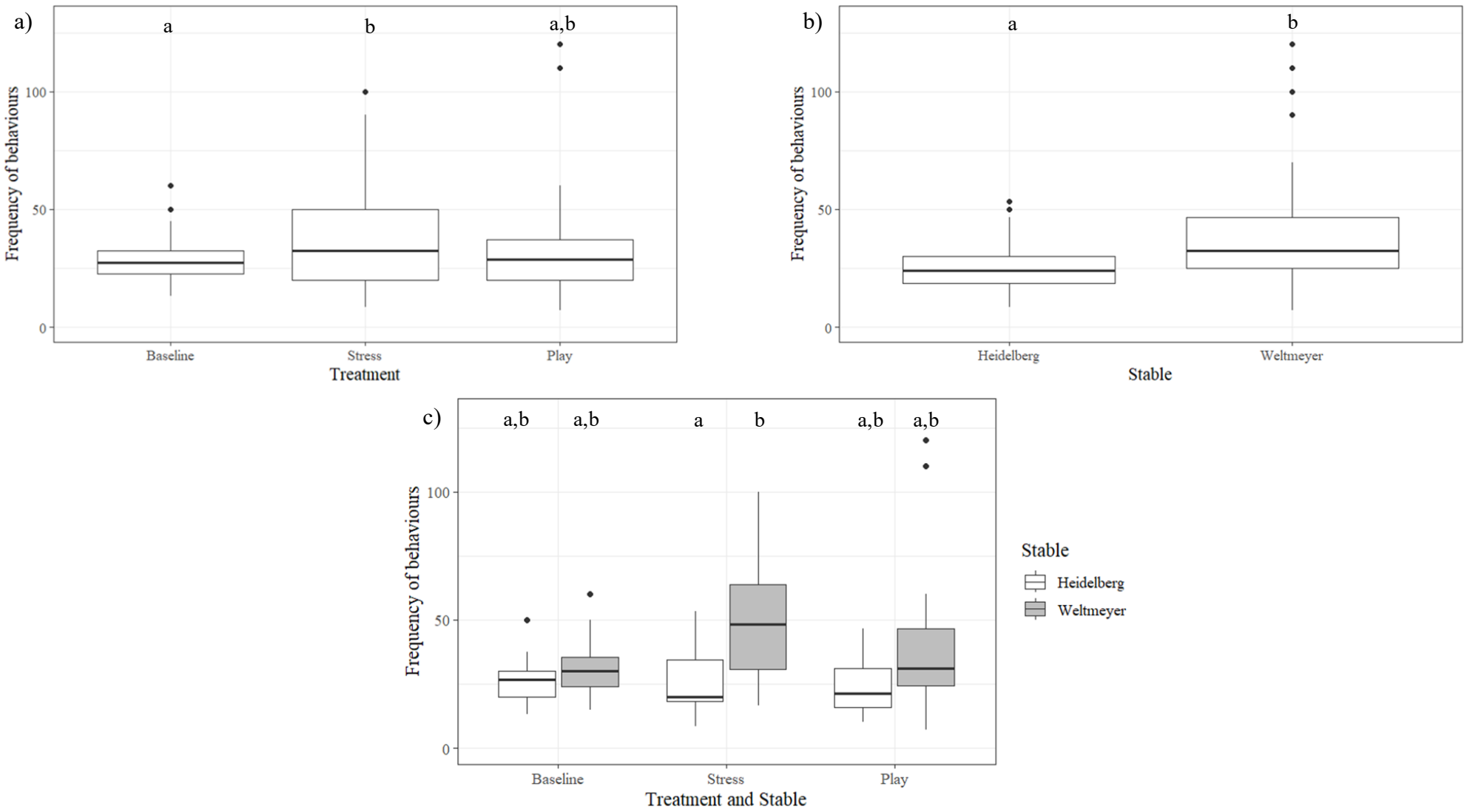


Figure 8. The frequency of behaviours before the decision point a) in three treatments, b) at both stables, and c) for the interaction between treatment x stable. Boxplots indicate the median (central dark line), the 1st and 3rd interquartiles (boxes), minimum and maximum scores (whiskers) and outliers (dots). Bars with the same alphabets are not significantly different.

Stable had a significant effect on the frequency of behaviours after the decision point (GLMM: $\chi^2 = 18.07$, df = 1, $p < 0.001$). Post hoc tests showed that the frequency of behaviours after the decision point of horses at the Weltmeyer stables (median = 38.75) was greater than those on the Heidelberg farm (median = 30.00) (Figure 9a). The interaction between treatment and stable was a significant predictor of the frequency of behaviours after the decision point (GLMM: $\chi^2 = 18.96$, df = 2, $p < 0.001$). Post hoc tests showed that the frequency of behaviours of the horses at the Weltmeyer stables was greater than those of the horses at the Heidelberg farm in the stress (median = 46.67 vs 23.75) and play treatments (median = 37.08 vs 27.50). There was no significant difference in the baseline treatment (median = 36.66 vs 31.25) (Figure 9b). Treatment was not a significant predictor of the frequency of behaviours after the decision point (GLMM: $\chi^2 = 5.11$, df = 2, $p = 0.078$). The median scores were as follows: baseline (33.33); stress (36.67); and play (30.00) treatments. There was no significant effect of sex on the frequency of behaviours after the decision point (GLMM: $\chi^2 = 0.02$, df = 1, $p = 0.885$). The median scores were as follows: females (34.17) and males (33.33). The interaction between treatment and sex was not a significant predictor of the frequency of behaviours after the decision point (GLMM: $\chi^2 = 0.66$, df = 2, $p = 0.719$). The median scores were as follows: baseline (females = 33.33 vs males = 33.33); stress (females = 36.66 vs males = 36.50); and play (females = 33.33 vs males = 30.00) treatments. Salivary cortisol was not a significant predictor of the frequency of behaviours before the decision point (GLMM: $\chi^2 = 0.36$, df = 2, $p = 0.546$).

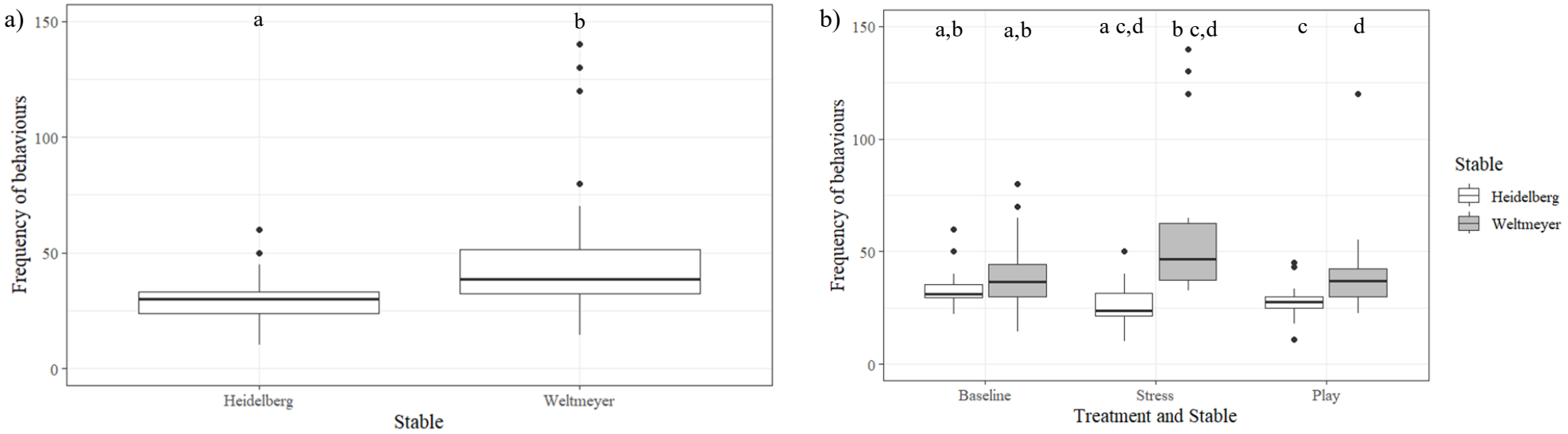


Figure 9. The frequency of behaviours after the decision point a) at both stables, b) showing the horses on both farms in the three treatments. Boxplots indicate the median (central dark line), the 1st and 3rd interquartiles (boxes), minimum and maximum scores (whiskers) and outliers (dots). Bars with the same alphabets are not significantly different.

A Spearman Rank regression showed a significant positive relationship ($R = 0.46$, $p < 0.001$) between the frequency of behaviours before and after the decision point in the maze, indicating that the frequency of behaviours of the horses was consistent when they entered and exited the maze (Figure 10).

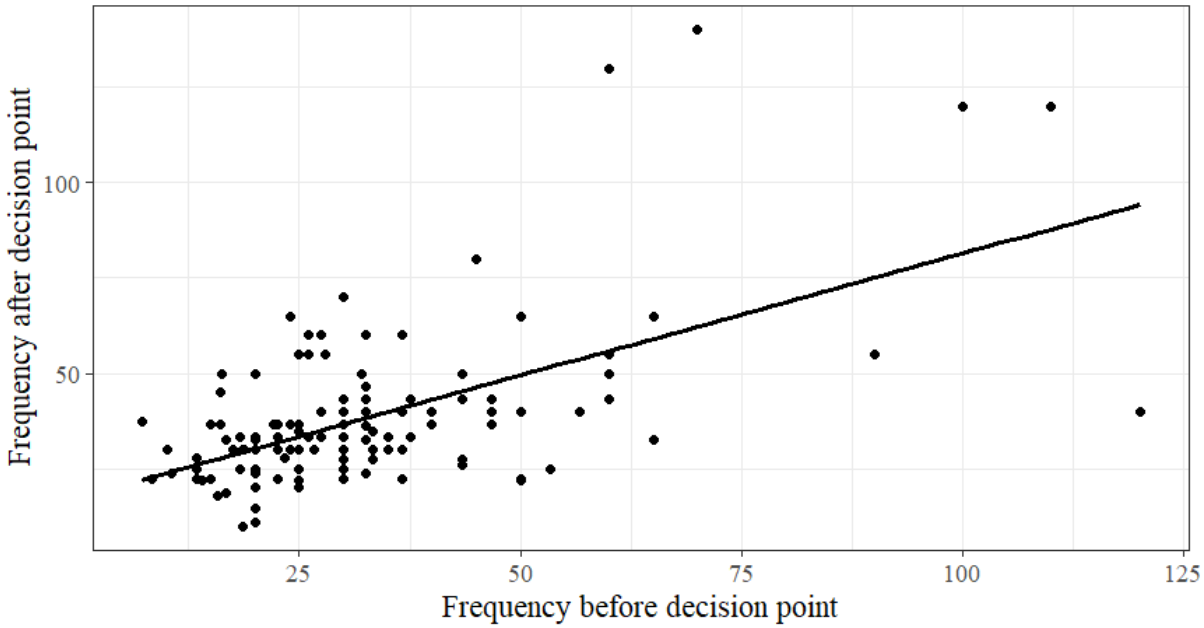


Figure 10. Regression of the frequency of behaviours before and after the decision point. The regression is significant ($p < 0.001$).

Latency

The latency was measured in seconds as the total time taken from the point when the horse entered the maze to the point where it exited the maze. From this, the total latency was divided into latency from the start of the maze to the decision point (latency before) and latency from the decision point to the end point of the maze (latency after).

Treatment was a significant predictor of latency before the decision point (GLMM: $\chi^2 = 8.15$, $df = 2$, $p = 0.017$). Post hoc tests showed that the latency in the stress treatment (median = 4.00) was faster than the latency in the play treatment (median = 4.00). There was no significant difference between the baseline (median = 4.00) and stress treatments and the baseline and play treatments for latency before the decision point (Figure 11a). Stable had a significant effect on the latency before the decision point (GLMM: $\chi^2 = 6.80$, $df = 1$, $p = 0.009$). Post hoc tests showed that the latency before the decision point of horses on the Heidelberg farm (median = 4.00) was slower than that of the horses at the Weltmeyer stables (median = 4.00) (Figure 11b). There was no significant effect of sex on the latency before the decision

point (GLMM: χ^2 = 2.98, df = 1, p = 0.084). The median scores were as follows: females (5.00) and males (4.00). The interaction between treatment and stable was not a significant predictor of the latency before the decision point (GLMM: χ^2 = 5.33, df = 2, p = 0.070). Median scores for baseline (Heidelberg = 4.00 vs Weltmeyer = 4.00); stress (Heidelberg = 4.00 vs Weltmeyer = 3.00); and play (Heidelberg = 4.00 vs Weltmeyer = 4.50) treatments. The interaction between treatment and sex was not a significant predictor of the latency before the decision point (GLMM: χ^2 = 1.53, df = 2, p = 0.466). The median scores were as follows: baseline (females = 5.00 vs males = 4.00); stress (females = 4.00 vs males = 3.50); and play (females = 5.00 vs males = 4.00) treatments. Salivary cortisol was not a significant predictor of the latency before the decision point (GLMM: χ^2 = 2.40, df = 2, p = 0.122).

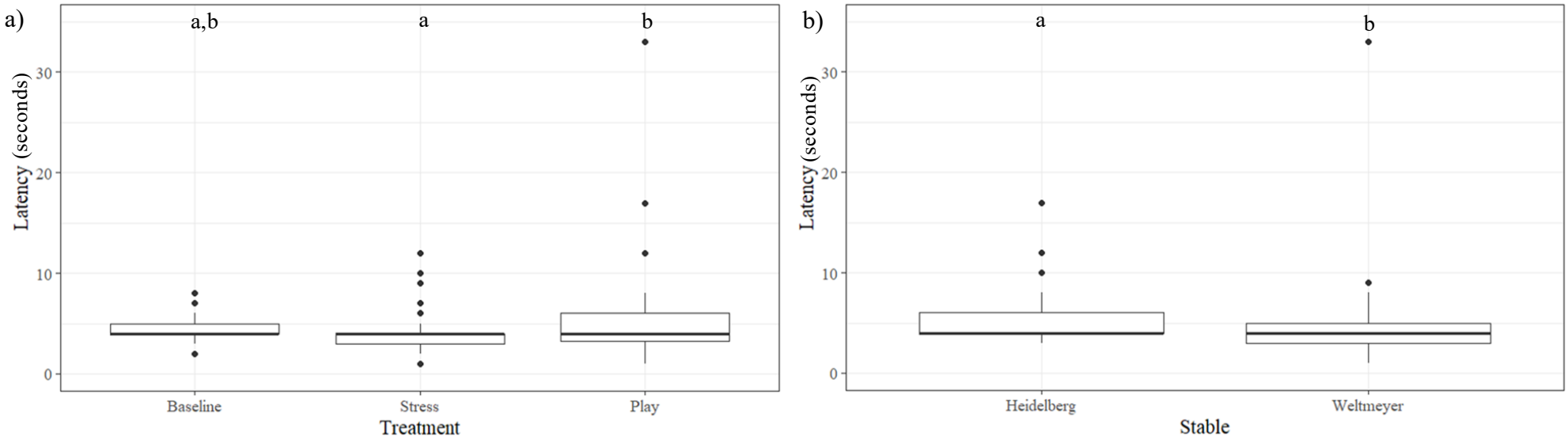


Figure 11. The latency before the decision point a) in three treatments, b) at both stables. Boxplots indicate the median (central dark line), the 1st and 3rd interquartiles (boxes), minimum and maximum scores (whiskers) and outliers (dots). Bars with the same alphabets are not significantly different.

Stable had a significant influence on the latency after the decision point (GLMM: χ^2 = 12.50, df = 1, $p < 0.001$). Post hoc tests showed that the latency after the decision point of horses at the Weltmeyer stables (median = 3) was faster than those on the Heidelberg farm (median = 4) (Figure 12a). The interaction between treatment and stable was a significant predictor of the latency after the decision point (GLMM: $\chi^2 = 14.40$, df = 2, $p < 0.001$). Post hoc tests showed that the latency of the horses at the Weltmeyer stables was faster than those at the Heidelberg farm in the stress (median = 4.00 vs 3.00) and play treatments (median = 4.00 vs 3.00). There was no significant difference in the baseline treatment (median = 3.00 vs 3.00) (Figure 12b). Treatment was not a significant predictor of latency after the decision point (GLMM: $\chi^2 = 1.31$, df = 2, $p = 0.519$). The median scores were as follows: baseline (3.00); stress (3.00); and play (3.00) treatments. There was no significant effect of sex on the latency after the decision point (GLMM: $\chi^2 = 0.33$, df = 1, $p = 0.565$). The median scores were as follows: females (3.00) and males (3.00). The interaction between treatment and sex was not a significant predictor of the latency after the decision point (GLMM: $\chi^2 = 2.33$, df = 2, $p = 0.312$). The median scores were as follows: baseline (females = 3.00 vs males = 3.00); stress (females = 3.00 vs males = 3.00); and play (females = 3.50 vs males = 3.00) treatments. Salivary cortisol was not a significant predictor of the latency after the decision point (GLMM: $\chi^2 = 0.01$, df = 2, $p = 0.940$).

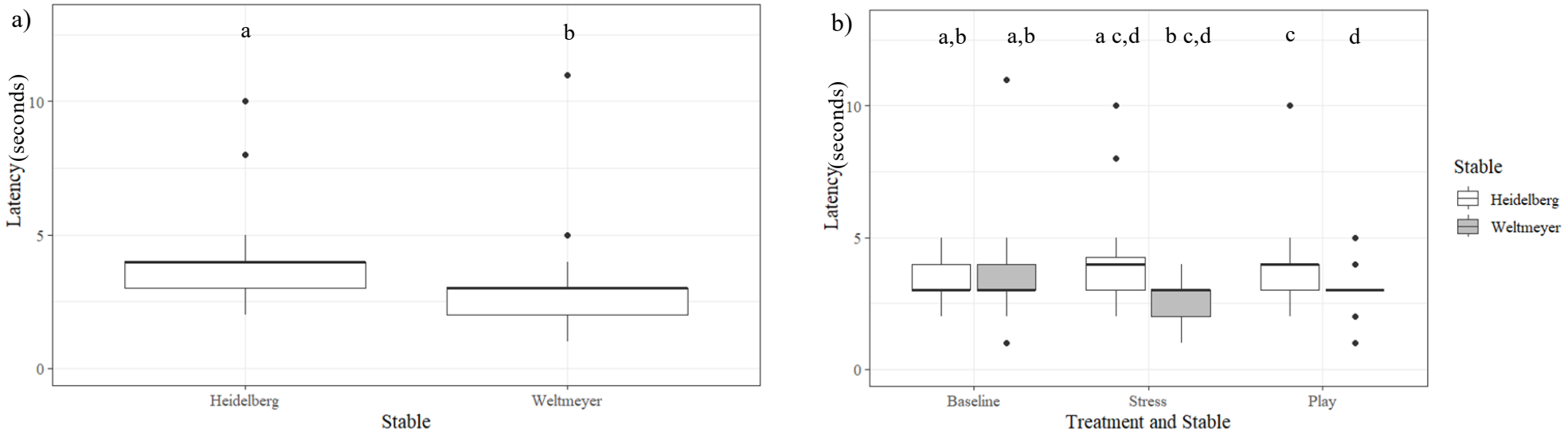


Figure 12. The latency after the decision point a) at both stables, b) showing the horses on both farms in the three treatments. Boxplots indicate the median (central dark line), the 1st and 3rd interquartiles (boxes), minimum and maximum scores (whiskers) and outliers (dots). Bars with the same alphabets are not significantly different.

A Spearman Rank regression showed a significant positive relationship ($R = 0.38$, $p < 0.001$) between the latency before and after the decision point in the maze, indicating that the latency of the horses was consistent when to the decision point and from the decision point to the exit of an arm (Figure 13).

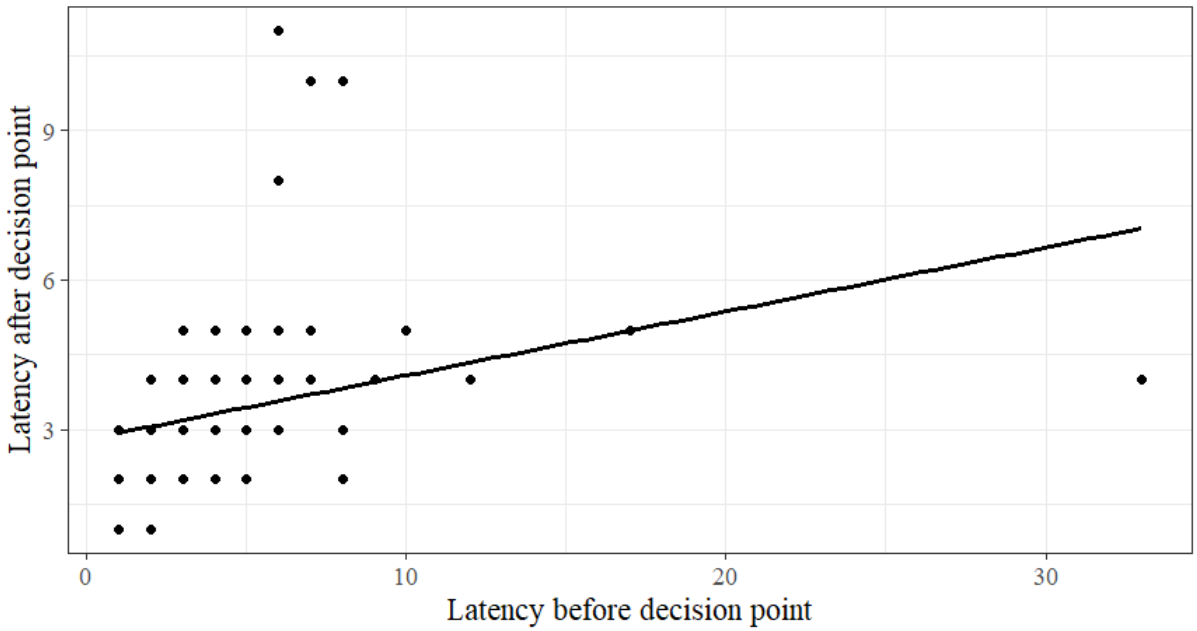


Figure 13. Regression of the latency before and after the decision point in the maze. The regression is significant ($p < 0.001$).

Emotional states

Emotional states were classified into positive, negative and neutral states (Table 2). Classification of a horse's emotional state was based on whether it showed 80% of the behaviours classified as positive, negative or neutral. Emotional states were not divided into before and after the decision point since very few differing behaviours occurred after the decision point.

Positive

Treatment was a significant predictor of positive emotional state (GLMM: $\chi^2 = 9.03$, df = 2, $p = 0.011$). Post hoc tests showed that the positive emotional state in the play treatment (median = 2.00) was greater than that in the stress treatment (median = 1.00). There was no significant difference between the baseline (median = 2.00) and stress treatments and the baseline and play treatments for positive emotional state in the maze (Figure 14). There was no significant effect observed of stable on positive emotional state (GLMM: $\chi^2 = 2.57$, df = 1, p =

0.109). The median scores were as follows: Heidelberg (1.00) and Weltmeyer (2.00). There was no significant effect of sex on positive emotional state (GLMM: $\chi^2 = 0.06$, df = 1, p = 0.805). The median scores were as follows: females (2.00) and males (2.00). The interaction between treatment and stable were not significant predictors of positive emotional state (GLMM: $\chi^2 = 0.74$, df = 2, p = 0.690). The median scores were as follows: baseline (Heidelberg = 2.00 vs Weltmeyer = 2.00); stress (Heidelberg = 1.00 vs Weltmeyer = 1.00); and play (Heidelberg = 2.00 vs Weltmeyer = 2.50) treatments. The interaction between treatment and sex was not a significant predictor of positive emotional state (GLMM: $\chi^2 = 2.87$, df = 2, p = 0.238). The median scores were as follows: baseline (females = 2.00 vs males = 2.00); stress (females = 1.00 vs males = 1.00); and play (females = 2.00 vs males = 2.00) treatments. Salivary cortisol was not a significant predictor of positive emotional state (GLMM: $\chi^2 = 0.45$, df = 1, p = 0.502).

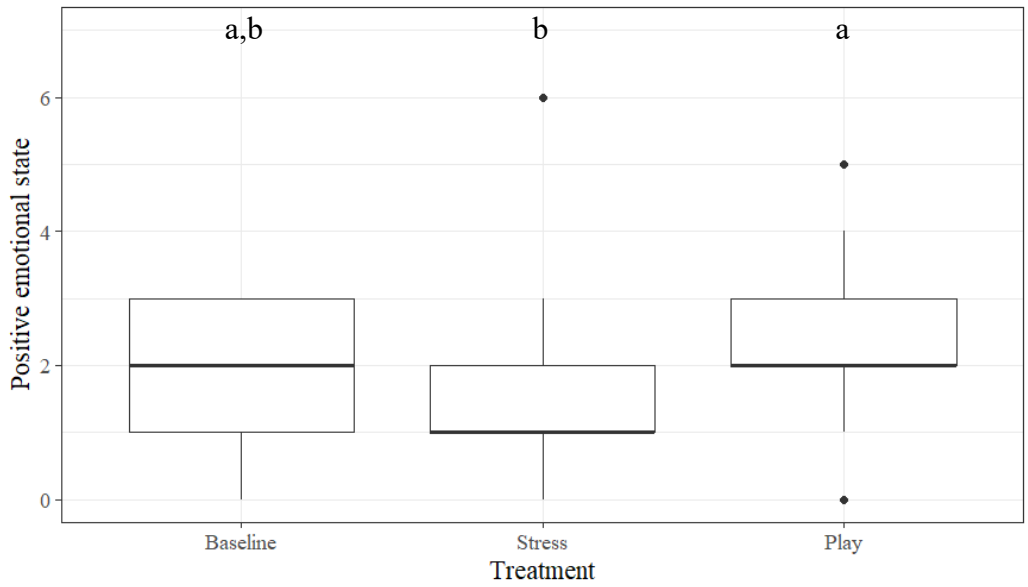


Figure 14. Positive emotional state in the three treatments. Boxplots indicate the median (central dark line), the 1st and 3rd interquartiles (boxes), minimum and maximum scores (whiskers) and outliers (dots). Bars with the same alphabets are not significantly different.

Negative

Treatment was a significant predictor of negative emotional state (GLMM: $\chi^2 = 20.15$, df = 2, p < 0.001). Post hoc tests showed that the negative emotional state in the stress treatment (median = 2.00) was greater than those in the baseline treatment (median = 1.00), and that the negative emotional state in the stress treatment was greater than those in the play treatment

(median = 1.00). There was no significant difference between the baseline and play treatments for negative emotional state (Figure 15a). A significant effect was observed of stable on negative emotional state (GLMM: $\chi^2 = 4.46$, df = 1, p = 0.035). Post hoc tests showed that the negative emotional state of the horses at the Weltmeyer stables (median = 2.00) was greater than those on the Heidelberg farm (median = 1.00) (Figure 15b). There was no significant effect of sex on negative emotional state (GLMM: $\chi^2 = 0.01$, df = 1, p = 0.942). The median scores were as follows: females (2.00) and males (1.50). The interaction between treatment and stable was not a significant predictor of negative emotional state (GLMM: $\chi^2 = 3.72$, df = 2, p = 0.155). The median scores were as follows: baseline (Heidelberg = 1.00 vs Weltmeyer = 1.00); stress (Heidelberg = 2.00 vs Weltmeyer = 2.50); and play (Heidelberg = 1.00 vs Weltmeyer = 2.00) treatments. The interaction between treatment and sex was not a significant predictor of negative emotional state in the maze (GLMM: $\chi^2 = 1.37$, df = 2, p = 0.503). The median scores were as follows: baseline (females = 1.00 vs males = 1.00); stress (females = 2.50 vs males = 1.00); and play (females = 2.00 vs males = 1.00) treatments. Salivary cortisol was not a significant predictor of negative emotional state (GLMM: $\chi^2 = 1.54$, df = 1, p = 0.214).

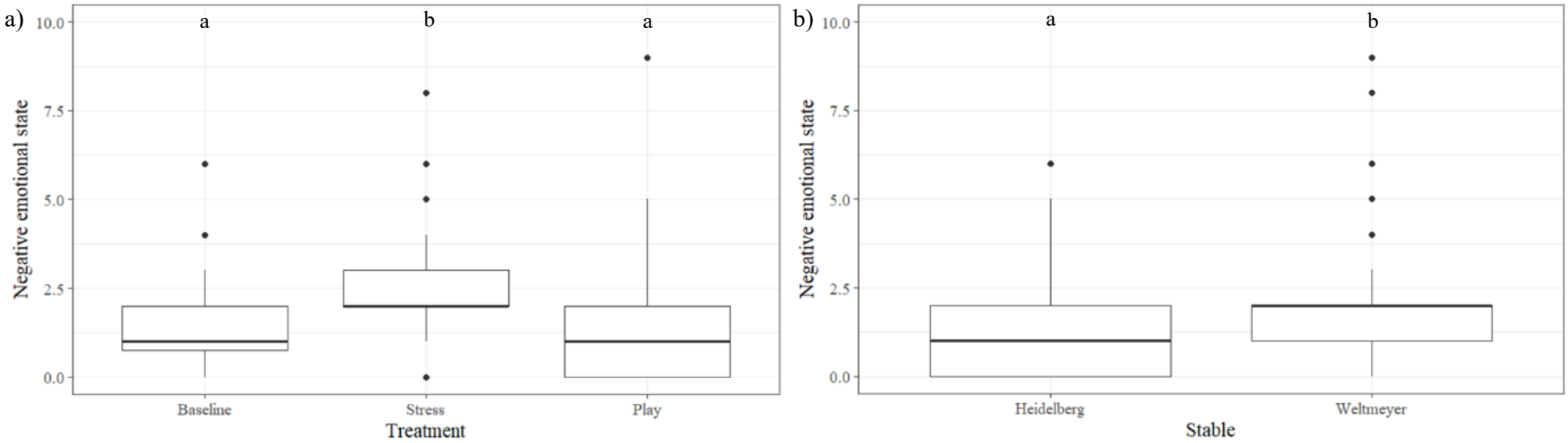


Figure 15. The negative emotional state a) in the three treatments, b) of the horses at both stables. Boxplots indicate the median (central dark line), the 1st and 3rd interquartiles (boxes), minimum and maximum scores (whiskers) and outliers (dots). Bars with the same alphabets are not significantly different.

Neutral

Treatment was not a significant predictor of neutral emotional state in the maze (GLMM: $\chi^2 = 0.65$, df = 2, p = 0.721). The median scores were as follows: baseline (10.00); stress (9.50); and play (9.00) treatments. There was no significant effect observed of stable on neutral emotional state (GLMM: $\chi^2 = 0.04$, df = 1, p = 0.846). The median scores were as follows: Heidelberg (9.00) and Weltmeyer (10.00). There was no significant effect of sex on neutral emotional state in the maze (GLMM: $\chi^2 = 1.79$, df = 1, p = 0.180). The median scores were as follows: females (10.00) and males (9.00). The interaction between treatment and stable was not a significant predictor of neutral emotional state in the maze (GLMM: $\chi^2 = 2.45$, df = 2, p = 0.294). The median scores were as follows: baseline (Heidelberg = 9.00 vs Weltmeyer = 10.00); stress (Heidelberg = 8.00 vs Weltmeyer = 10.50); and play (Heidelberg = 9.00 vs Weltmeyer = 8.50) treatments. The interaction between treatment and sex was not a significant predictor of neutral emotional state in the maze (GLMM: $\chi^2 = 0.20$, df = 2, p = 0.903). The median scores were as follows: baseline (females = 10.00 vs males = 9.00); stress (females = 11.00 vs males = 9.00); and play (females = 9.00 vs males = 9.00) treatments. Salivary cortisol was not a significant predictor of neutral emotional state in the maze (GLMM: $\chi^2 = 1.50$, df = 1, p = 0.220).

Initial decision success and final decision success

An initial decision was recorded as successful (Y) when the arm of the maze with the preferred food type was selected or unsuccessful (N) when the arm of the maze with the rocks in the feeding bowl was selected, depending on the position of the horse's head at the decision point. A final decision was recorded as successful (Y) when the arm of the maze with the preferred food type was selected or unsuccessful (N) when the arm of the maze with the rocks in the feeding bowl was selected. Overall, on 66 (49%) occasions, the horses made the correct initial decision up to the decision point and also made the correct final decision for the food at the end of an arm (Y-Y). On 45 (33%) occasions, the horses were unsuccessful in their initial decision and the final decision (N-N). On 14 (10%) occasions, the horses did not make the correct initial decision, but then self-corrected and made the correct final decision (N-Y). Finally, on 11 (8%) occasions, the horses made the correct initial decision but then made the incorrect final decision (Y-N) (Figure 16a). A Fisher's exact test showed a significant

association between the overall initial and final decisions ($p < 0.001$). The odds of a Y-Y response were 1.9 times more common than a Y-N response (Figure 16a).

Based on the above scenarios, in the baseline treatment (Figure 16b), Y-Y decisions were made on 26 (38%) occasions, N-N decisions on 27 (40%) occasions, N-Y on 6 (9%) occasions and Y-N on 9 (13%) occasions. A Fisher's exact test showed a non-significant association between the overall initial and final decisions ($p = 0.563$). In the stress treatment (Figure 16c), Y-Y decisions were made on 19 (56%) occasions, N-N decisions on 11 (32%) occasions, N-Y decisions on 3 (9%) occasions and Y-N on 1 (3%) occasion. A Fisher's exact test showed a non-significant association between the overall initial and final decisions ($p = 0.283$). In the play treatment (Figure 14d), Y-Y decisions were made on 21 (62%) occasions, N-N decisions on 7 (20%) occasions, N-Y decisions on 5 (15%) occasions and Y-N decisions on 1 (3%) occasion. A Fisher's exact test showed a significant association between the overall initial and final decisions ($p = 0.014$) The odds of a Y-Y response were 13.7 times more common than a Y-N response (Figure 16d).

a)

Final \ Initial	Yes	No
Yes	66	14
No	11	45

b)

Final \ Initial	Yes	No
Yes	26	6
No	9	27

c)

Final \ Initial	Yes	No
Yes	19	3
No	1	11

d)

Final \ Initial	Yes	No
Yes	21	5
No	1	7

Figure 16: Matrices showing the initial and final decisions counts for a) all of the horses in all treatments, b) baseline treatment, c) stress treatment and d) play treatment.

The horses on the Heidelberg farm made the following decisions (Figure 17a): Y-Y decisions were made on 29 (45%) occasions, N-N decisions on 23 (36%) occasions, N-Y on 3 (5%) occasions and Y-N on 9 (14%) occasions. A Fisher's exact test showed a non-significant association between the overall initial and final decisions for the horses on the Heidelberg farm ($p = 0.331$). The horses at the Weltmeyer stables made the following decisions (Figure 17b):

Y-Y decisions were made on 37 (51%) occasions, N-N decisions on 22 (31%) occasions, N-Y decisions on 8 (11%) occasions and Y-N on 5 (7%) occasions. A Fisher's exact test showed a non-significant association between the overall initial and final decisions of the horses at the Weltmeyer stables (p = 0.130). The female horses made the following decisions (Figure 17c): Y-Y decisions were made on 20 (50%) occasions, N-N decisions on 11 (27%) occasions, N-Y decisions on 4 (10%) occasions and Y-N decisions on 5 (13%) occasions. A Fisher's exact test showed a non-significant association between the overall initial and final decisions for female horses (p = 0.705). The male horses made the following decisions (Figure 17d): Y-Y decisions were made on 46 (48%) occasions, N-N decisions on 34 (36%) occasions, N-Y decisions on 7 (7%) occasions and Y-N decisions on 9 (9%) occasions. A Fisher's exact test showed a non-significant association between the overall initial and final decisions for male horses (p = 1.000).

a)

Final \ Initial	Yes	No
Yes	29	3
No	9	23

b)

Final \ Initial	Yes	No
Yes	37	8
No	5	22

c)

Final \ Initial	Yes	No
Yes	20	4
No	5	11

d)

Final \ Initial	Yes	No
Yes	46	7
No	9	34

Figure 17: Matrices showing the initial and final decisions counts for a) horses on the Heidelberg farm, b) horses at the Weltmeyer stables, c) all female horses and d) all male horses.

Initial decision

Treatment was not a significant predictor of the proportion of initial decision success (GLMM: $\chi^2 = 1.82$, df = 2, p = 0.403). The proportion of successful and non-successful tests in the baseline treatment was 0.515 and 0.485, in the stress treatment was 0.588 and 0.412, and in the play treatment was 0.647 and 0.353. The success of a test was greatest in the play treatment. The horses at the Weltmeyer stables were more successful than those at the

Heidelberg farm. However, stable was not a significant predictor of the proportion of initial decision success (GLMM: $\chi^2 = 1.48$, df = 1, p = 0.224). The proportion of successful and non-successful tests at the Heidelberg farm was 0.5 and 0.5 and was 0.625 and 0.375 at the Weltmeyer stables. The female horses were more successful than the male horses. However sex was not a significant predictor of the proportion of initial decision success (GLMM: $\chi^2 = 0.18$, df = 1, p = 0.674). The proportion of successful and non-successful tests of female horses was 0.6 and 0.4 and was 0.552 and 0.448 for the male horses. Overall the horses at the Weltmeyer stables had a greater success than the horses at the Heidelberg farm across all treatments. However, the interaction between treatment and stable was not a significant predictor of the proportion of initial decision success (GLMM: $\chi^2 = 0.28$, df = 2, p = 0.870). The proportion of successful and non-successful tests in the baseline treatment was 0.469 and 0.531 for the horses on the Heidelberg farm and 0.556 and 0.444 for the horses at the Weltmeyer stables. The proportion of successful and non-successful tests in the stress treatment was 0.5 and 0.5 for the horses on the Heidelberg farm and 0.667 and 0.333 for the horses at the Weltmeyer stables. The proportion of successful and non-successful tests in the play treatment was 0.562 and 0.438 for the horses on the Heidelberg farm and 0.722 and 0.278 for the horses at the Weltmeyer stables. Overall the male horses had greater success in the stress and play treatments while the female horses had greater success in the baseline treatments. However, the interaction between treatment and sex was not a significant predictor of the proportion of initial decision success (GLMM: $\chi^2 = 2.53$, df = 2, p = 0.282). The proportion of successful and non-successful tests in the baseline treatment was 0.65 and 0.35 for the female horses and 0.458 and 0.542 for the male horses. The proportion of successful and non-successful tests in the stress treatment was 0.5 and 0.5 for the female horses and 0.625 and 0.375 for the male horses. The proportion of successful and non-successful tests in the play treatment was 0.6 and 0.4 for the female horses and 0.667 and 0.333 for the male horses. Salivary cortisol was not a significant predictor of the proportion of initial decision success (GLMM: $\chi^2 = 1.81$, df = 1, p = 0.179).

Final decision

Treatment was a significant predictor of the proportion of final decision success (GLMM: $\chi^2 = 7.92$, df = 2, p = 0.019) (Figure 18). The proportion of successful and non-successful tests in the baseline treatment was 0.471 and 0.529, in the stress treatment was 0.647 and 0.353, and in the play treatment was 0.765 and 0.235. The success of a test was greatest in

the play treatment. Stable was not a significant predictor of the proportion of final decision success (GLMM: $\chi^2 = 0.09$, df = 1, p = 0.768). The proportion of successful and non-successful tests at the Heidelberg farm was 0.594 and 0.406 and was 0.583 and 0.417 at the Weltmeyer stables. The horses at both farms had similar success rates. The female horses were more successful than the male horses, but sex was not a significant predictor of the proportion of final decision success (GLMM: $\chi^2 = 0.71$, df = 1, p = 0.399). The proportion of successful and non-successful tests of female horses was 0.625 and 0.375 and was 0.573 and 0.427 for the male horses. Overall, the horses at the Weltmeyer stables had greater success than the horses at the Heidelberg farm across in the stress and play treatments, but the horses on the Heidelberg farm had greater success in the baseline treatment compared to the horses at the Weltmeyer stables. However, the interaction between treatment and stable was not a significant predictor of the proportion of final decision success (GLMM: $\chi^2 = 2.11$, df = 2, p = 0.349). The proportion of successful and non-successful tests in the baseline treatment was 0.531 and 0.469 for the horses on the Heidelberg farm and 0.417 and 0.583 for the horses at the Weltmeyer stables. The proportion of successful and non-successful tests in the stress treatment was 0.562 and 0.438 for the horses on the Heidelberg farm and 0.722 and 0.278 for the horses at the Weltmeyer stables. The proportion of successful and non-successful tests in the play treatment was 0.75 and 0.25 for the horses on the Heidelberg farm and 0.778 and 0.222 for the horses at the Weltmeyer stables. Overall the male horses had greater success in the stress treatment while the female horses had greater success in the baseline and play treatments. However, the interaction between treatment and sex was not a significant predictor of the proportion of final decision success (GLMM: $\chi^2 = 0.69$, df = 2, p = 0.707). The proportion of successful and non-successful tests in the baseline treatment was 0.55 and 0.45 for the female horses and 0.438 and 0.562 for the male horses. The proportion of successful and non-successful tests in the stress treatment was 0.6 and 0.4 for the female horses and 0.667 and 0.333 for the male horses. The proportion of successful and non-successful tests in the play treatment was 0.8 and 0.2 for the female horses and 0.75 and 0.25 for the male horses. Salivary cortisol was not a significant predictor of the proportion final decision success (GLMM: $\chi^2 = 0.50$, df = 1, p = 0.480).

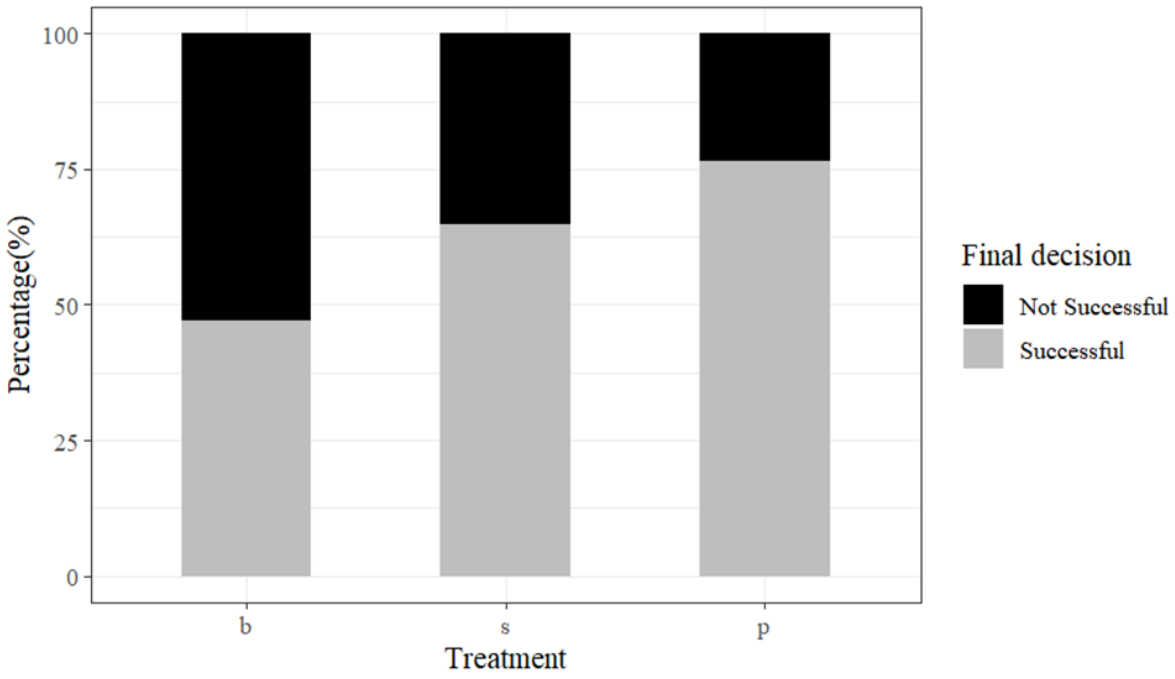


Figure 18. The final decision of the horses in the three treatments.

Salivary cortisol

Salivary cortisol was sampled from the horses in each treatment. The saliva samples were then tested, using the LS-MS/MS method, by Lancet Laboratories to determine the level of cortisol (nmol/l) in the saliva samples. The results follow below (Table 3).

Treatment was not a significant predictor of salivary cortisol (GLMM: $\chi^2 = 3.28$, df = 2, p = 0.194). The median scores were as follows: baseline (1.25); stress (1.20); and play (1.30) treatments. There was no significant effect of stable on salivary cortisol (GLMM: $\chi^2 = 0.02$, df = 1, p = 0.895). The median scores were as follows: Heidelberg (1.20) and Weltmeyer (1.25). There was no significant effect of sex on salivary cortisol (GLMM: $\chi^2 = 1.20$, df = 1, p = 0.273). The median scores were as follows: females (1.10) and males (1.30). The interaction between treatment and stable was not a significant predictor of salivary cortisol (GLMM: $\chi^2 = 3.02$, df = 2, p = 0.221). The median scores were as follows: baseline (Heidelberg = 1.30 vs Weltmeyer = 1.20); stress (Heidelberg = 1.20 vs Weltmeyer = 1.20); and play (Heidelberg = 1.15 vs Weltmeyer = 1.40) treatments. The interaction between treatment and sex was not a significant predictor of salivary cortisol (GLMM: $\chi^2 = 0.51$, df = 2, p = 0.774). The median scores were as follows: baseline (females = 1.10 vs males = 1.30); stress (females = 1.15 vs males = 1.20); and play (females = 1.10 vs males = 1.35) treatments.

Table 3: Salivary cortisol (nmol/l) change during experimental treatments. Part 1: the first round of experiments. Part 2: the repeated experiments.

Horse	Baseline stress (nmol/l)		Stress (nmol/l)		Baseline play (nmol/l)		Play (nmol/l)	
	Part 1	Part 2	Part 1	Part 2	Part 1	Part 2	Part 1	Part 2
1	1.5	2.3	1.8	1.9	2.1	1.5	1.7	2.8
2	1.5	1.1	1	0.9	1.7	1	0.8	1.3
3	1.8	1.4	1.6	0.8	1.8	1.6	1.2	1.6
4	0.9	0.8	0.9	1.2	0.9	1	1	0.8
5	2.2	1.8	1.5	1	1	1.7	1.1	1.1
6	1.8	1	1.1	1.2	1.2	1.2	1.1	0.9
7	1.3	1.5	1.1	1.4	1.2	0.9	1.5	1.3
8	0.9	0.9	1.1	1.2	1.8	1.3	2	0.9
9	0.8	1.4	0.8	1	1.9	0.9	1.9	1.2
10	1.1	1.1	1.4	1	1.4	1.2	1.5	0.9
11	1.6	0.8	1.2	0.9	2.2	1	1.5	0.8
12	1	0.9	1.3	1.1	1.1	1.1	1.2	1
13	1	1	0.9	1.4	0.9	1.3	1.3	1.4
14	1.3	1	1.2	1	1.1	1.2	1.8	2
15	2	1.3	2	1.8	1.6	1.3	1.4	1.5
16	1	1.5	1.3	1.1	1.5	2.8	1.1	2.5
17	1.7	1.2	1.3	1.4	1.6	1.5	1.2	2.1

DISCUSSION

The aim of my study was to ascertain whether and how play behaviour can ameliorate acute stress in horses (*Equus caballus*) from two different housing and working conditions. My study involved recording the behaviours of individual horses exposed to three treatments (baseline, stress and play). I also measured the salivary cortisol levels of each horse before and after their exposure to the treatments.

General behaviours

Behavioural diversity was recorded as the number of unique behaviours in the maze. The diversity of behaviours before the decision point in the stress treatment was significantly greater than those in the baseline treatment Behavioural diversity refers to the range and variability of behaviours shown by individuals within a population. The levels of behavioural diversity can have conflicting behavioural motivations, leading to an ongoing debate as to whether behavioural diversity could potentially be a positive indicator of animal welfare (Miller *et al.* 2020). On the one hand, low behavioural diversity is an indicator of animal boredom or that they are merely in a rhythm of habit (Rabin 2003; Miller *et al.* 2020). On the other hand, higher behavioural diversity indicates that an animal has adapted well to a stimulating or enriched environment, and so it has been considered an indicator of good animal welfare (Miller *et al.* 2020, Hall *et al.* 2021). For example, the higher the diversity of behaviours in bottlenose dolphins (*Tursiops truncatus and Tursiops aduncus*), the lower their faecal cortisol levels (Miller *et al.* 2021), indicating reduced stress. Similarly, chimpanzees (*Pan troglodytes*) also showed increased behavioural diversity when they were in a stimulating or enriched environment, specifically a captive environment, which was inversely related to their faecal cortisol levels (Hall *et al.* 2021).

Despite the absence of corroborating physiological support (cortisol levels were similar regardless of treatment), I still maintain that the high behavioural diversity of the horses indicates uncertainty and stress. When animals are exposed to new or stressful conditions, they may exhibit a wider range of behaviours to cope with the challenges. For example, Koolhass *et al*. (1999), found that male rats (species unknown) displayed a wider variety of aggressive behaviours when exposed to a prod that provided the rats with a mild shock when touched. Réale et al. (2007) found that some animals behaved more boldly or more aggressively, displaying a wide range of behaviours in stressful situations. Based on these studies, the greater behavioural diversity in the stress treatment in my study, indicates that the horses experienced

a negative situation or stimulus. This is in keeping with the theory that animals display coping behaviours, like quick movements, head jerking and freezing, to name a few, in unfamiliar or stressful situations (Wechsler 1995; Hall *et al.* 2018; Anderson *et al.* 2020).

Based on my initial assumption of greater behavioural diversity indicating a coping response to stress, the reduction of behavioural diversity in the play treatment, matching the baseline treatment, supports my prediction that play ameliorates stress. In other words, play is effectively a coping mechanism following a stressor (see below).

The frequency of behaviours was recorded as the number of times each behaviour occurred per 10 seconds in the maze. The frequency of behaviours before the decision point in the stress treatment was significantly greater than those in the baseline treatment (Figure 8a). The frequency of behaviours of an animal is closely linked to the diversity of behaviours (Wemelsfelder *et al.* 2000). In the same way that increased behavioural diversity in a stressful situation has a negative connotation, a greater frequency of behaviours also has a negative connotation (Verbruggen and Logan 2008). Although Verbruggen and Logan (2008) do not directly consider stressful situations, they explain that a lack of control or impulsivity can be a result of increased frequency of behaviours which can lead to negative consequences in some instances. An increase in frequency of behaviours can also have energetic costs as seen in the optimal foraging theory, where animals minimize their behaviours and movements while still obtaining enough food (Pyke and Starr 2021). Thus, a greater investment in different and more behaviours implies a greater energetic cost which is not optimal. The horses displayed a significant increase in the frequency of behaviours in the stress treatment compared to the baseline treatment, indicating that they experienced a negative situation or stimulus in the stress treatment. The frequency of behaviours was similar in the baseline and play treatments.

The positive association of both behavioural diversity and frequency and the period before and after the decision point in the maze indicates that the horses did not change their responses within the maze. Behavioural consistency in a maze often reflects sequential temporal carry over effects (i.e. initial behaviours are repeated later). For example, Wistar male rats which were exposed to an elevated plus-maze continuously displayed a similar array of behaviours within the first 24 hours in the maze (Andrade *et al.* 2003).

Latency was a measure of the time it took the horse to enter and then exit the maze. The horses moved to the decision point significantly faster in the stress treatment compared to the play treatment. Animals respond to stress in many different ways. Some animals may slow

down or speed up their movements and behaviours in a stressful situation (Baragli *et al*. 2014; Ellis *et al*. 2014; Esch *et al.* 2019). Metz *et al.* (2001) showed that laboratory rats had a slower latency to enter a compartment when they were stressed. In contrast, Boonstra (2004) found that when stressed, Arctic ground squirrels (*Urocitellus parryii*) sped up their movements to avoid predators (Boonstra 2004). In my study, the horses responded more quickly in the stress treatment, possibly to distance themselves from a stressful situation, even though they could exit the enclosure. The horses did not move as quickly through the play treatment, potentially because play ameliorated the stress response.

The emotional states were classified into positive, negative and neutral states according to the diversity and frequency of behaviours observed. A positive emotional state was significantly greater in the play treatment compared to that of the stress treatment, whereas a negative emotional state was significantly more prevalent in the stress treatment compared to that in the baseline and play treatments. Similarly, domesticated pigs (*Sus scrofa*) showed a positive emotional state associated with play behaviour and a negative emotional state associated with a stressful or unfamiliar environment (Murphy *et al*. 2014). Lowland or spotted paca (*Cuniculus paca*) have also been shown to display a positive emotional state when interacting with an object in object play (Lima *et al.* 2023).

Taken together, my data showed that, using several behavioural metrics, the horses displayed heightened responses to the stressor. During the stress treatment, the horses displayed behaviours that could be considered as fearful or anxious (skittish movements, faster movements, freezing, wide eyes, head raised on alert). They most likely viewed the auditory stressor as a threat (Sapolsky 2002; Bartolomé and Cockram 2016). When a stimulus is perceived as a threat, a neuroendocrine response occurs via the Hypothalamic Pituitary Adrenal (HPA) axis which causes the release of adrenocorticotrophic hormone (ACTH) (Kang and Lee 2016). This stimulates the release of cortisol from the adrenal cortex into the blood stream (Koolhaas *et al*. 2011; Bartolomé and Cockram 2016; Kang and Lee 2016). Such a stress response is evident in the stress treatment (but see below about cortisol). In contrast, the responses in the play treatment were similar to the baseline treatment, when there was no auditory stressor.

Play behaviours can enhance a stress response, as seen during social play in horses which experienced chronic stress (Hausberger *et al.* 2012). In contrast, play can reduce the stress response as seen in laboratory rats (*Rattus norvegicus*) that were play-fighting and did

not display any forms of stress (Pellis and Pellis 2017). Neurobiologically, play is able to activate certain brain signalling pathways, one of which is the transmission of norepinephrine, the neurotransmitter of epinephrine, which plays a role in stress signalling pathways, which ameliorate the effect of a stressor (Wang and Aamodt 2012). An example from DeVries *et al.* (2003) showed that interactions (like play) with other animals can modulate the HPA axis in animals in a way that reduces their levels of stress.

Decision-making

I tested the cognition of the horses by assessing their decision-making in the maze. In the play treatment, the horses made the correct initial and final decisions 13.7 times more often than the incorrect final decision. These odds were significantly better than in the baseline and stress treatments.

Stress can influence the decision-making of animals from a physiological perspective. There is a link between the dysregulation of the HPA axis and cognitive bias as seen in merino sheep (*Ovis aries*), where prolonged periods of chronic stress led to the dysregulation of the HPA axis (Verbeek *et al.* 2019). Stress also influences decision-making of animals which make rash and incorrect or potentially harmful decisions (Starcke and Brand 2012). For example, prey species run from a predator as opposed to hiding or freezing in order to avoid a predator that is potentially faster or more agile than the prey species (Barrett 2015). This is partially supported by the faster latency in the stress treatment, indicating that quicker movement through the maze, resulted in rash decisions. However, this association did not hold true for the baseline treatment, where latency was longer, but decision-making was poorer. The greatest success in decision-making by the horses was in the play treatment where there was a significant association between the initial and final decisions. The reason for this success could be explained by the Yerkes-Dodson law which states that an animal will perform at its best cognitively at an optimal level of stress (Yerkes-Dodson, 1908). This law could also be used to explain why the horses did not perform optimally during the baseline treatment because they were not experiencing any form of stress at that point (Yerkes-Dodson, 1908). During the play treatment, the horses were exposed to a stressor initially and then allowed to interact and play with a ball. This means that while the horses showed similar behavioural metrics in the play and baseline treatments (see above), they were better able to make decisions. Whether they were in an optimal stress state needs further study.

Stable conditions

The horses at the Weltmeyer stables displayed significantly greater behavioural diversity and frequency before and after the decision point than the horses at the Heidelberg farm. In addition, the latency of Weltmeyer horses was faster than that of the Heidelberg horses before and after the decision point. Weltmeyer horses were in a comparatively negative emotional state than the Heidelberg horses. Taken together, the Weltmeyer horses appeared to be more stressed than the horses at the Heidelberg farm. Yarnell *et al.* (2015) demonstrated that horses housed under conditions that reflected their conditions in nature were less stressed than horses that lived in more constrained conditions, apparently because they could not interact with other horses as they would in nature. In another example, laboratory rats housed under constrained conditions with no enrichment and limited space were more stressed than rats not housed under these conditions (Sherwin 2004). Similarly, the horses stabled in this study were more affected by the stressor than the horses on the more natural Heidelberg farm.

One explanation for the differences between the horse groups could be that the stabled horses are ridden more regularly than the horses on the Heidelberg farm. Therefore, the Weltmeyer horses potentially experienced chronic stress, which can be directly related to how often the horses are being ridden. Hausberger *et al.* (2012) found that horses that were ridden each day by multiple riders with different skill sets experienced chronic stress due to problems encountered in their vertebral columns. Another explanation could be that the horses at the Weltmeyer stables spent half of the day in a confined space (stable) if they were not ridden on a particular day, resulting in boredom, bad habits and overall increased stress due to lack of care or activity (Hausberger *et al.* 2012; Yarnell *et al.* 2015; Esch *et al.* 2019). However, there was no evidence to suggest that the horses under the less constrained conditions at the Heidelberg farm made better decisions compared to the horses under more constrained conditions. This indicates that decision-making was not affected by stabling conditions.

Sex effects

Only one sex effect occurred in my study. Male horses showed a greater diversity of behaviours before the decision point in the maze compared to the female horses. Based on the above assumptions that the greater the diversity of behaviours indicate greater stress, this would mean that the male horses were more stressed than the female horses. This is contrary to the literature. Generally, females display more behaviours associated with stress, like being skittish

or more fearful, as seen in various predator-prey examples (Barrett 2015; Han *et al.* 2015; Donelan and Trussel 2020). For example, Han, *et al.* (2015) showed that male common water striders (*Gerris gracilicornis*) were more bold in the presence of predators compared to the female water striders, which showed more fearful behaviours linked to stress. In contrast, there are instances where female animals display bolder behaviours in stressful situations. For example, female guppies (*Poecilia reticulata*) tend to display bolder behaviours in the presence of male guppies due to past experiences where the male guppies sexually harassed the females (Piyapong *et al.* 2009). Another example is seen where male field crickets (*Teleogryllus commodus)* have an increased metabolic rate in the presence of a predator compared to females, which is directly linked to the release of stress hormones into the body (Lagos and Herberstein 2017), thus confirming that the male field crickets responded in a more stress-like manner when confronted by a predator. There is no evidence that male horses made better decisions compared to the female horses.

Salivary cortisol

Neither treatment, stable nor sex were significant predictors of salivary cortisol in the horses. In addition, salivary cortisol was not a significant covariate in any of the analyses. It appears that the horses were not in a physiologically stressed state during the stress treatment, which is in contrast to the behavioural metrics. It has been suggested that behavioural and physiological responses to a stressor are not always coupled but occur independently of one another (Hall *et al.* 2018; Safryghin *et al.* 2019). In particular, Hall *et al.* (2018) suggested that changes in salivary cortisol in horses provide an indication of long term emotional state or mood rather than in the short term when the horse reacts to a stressor, as in my study. If so, it could be that the changes in salivary cortisol could not be detected in my study. In another study, salivary cortisol levels in horses were similar before and after a cognitive task, possibly because horses were not in a stressed state (Olczak *et al.* 2018).

Another explanation for the lack of a treatment effect on salivary cortisol could be the sampling interval between stressor and the collection of saliva. The sampling interval between the stressor and collection of the saliva for cortisol (i.e. about 10 minutes) might have been too short to detect cortisol changes. I based the timing of saliva sampling on a study that sampled saliva from horses 10 minutes before and after they underwent a food motivation task (Olczak *et al.* 2018). In contrast to this, a recent study, Holtby *et al.* (2023), suggested that a saliva

sample be taken from horses up to 30 minutes after their exposure to a mild stressor or activity. I was unaware of the findings of this study when I conducted my research.

CONCLUSION

I studied whether and how play behaviours can ameliorate acute stress in horses (*Equus caballus*) from two different stabling and working conditions. It was evident that the horses exhibited behaviours associated with positive emotional state and made better decisions when they interacted or played with an object (ball) after being exposed to a stressor. This was in contrast to the stress treatment in which the horses were stressed. The similarity between the play and baseline treatment for the general behaviours suggest that play returned the horses to a baseline state. Nonetheless, decision-making did not show a similar pattern while play actually enhanced cognition.

There are mixed reports about the relationship between play and stress. In Sprague-Dawley rats play ameliorated stress and produced a sense of resilience to stress in the rats (Burgdorf *et al.* 2017). In another example, Lonsdorf *et al.,* (2014) found lower cortisol levels in faecal samples of chimpanzees when they were exhibiting play behaviour. While the exact relationship between stress and play is still unknown, it is evident that play behaviour influences stress in animals in some positive way. It is because of this that many researchers have debated whether play behaviour could be used as a measure of animal welfare (Miller *et al*. 2020). In contrast to these findings, play behaviour was present in riding school horses that displayed signs of chronic stress (Hausberger *et al*. 2012) and play behaviour increased when rhesus monkeys (*Macaca mulatta*) displayed social stress (Loy 1970; Held and Špinka 2011). These studies indicate that play behaviour could not be used as an indicator of positive animal welfare. However, other studies, like those mentioned above, found a clear link between positive emotional state and play behaviour (Baragli *et al*. 2014; Murphy *et al*. 2014). For example, play behaviour in domesticated pigs (*Sus scrofa*) was directly linked to positive emotional states (Murphy *et al*. 2014). In other studies, it was evident that play behaviour had long term benefits for animals in adapting to novel environments (Spinka *et al*. 2001; Pellis and Pellis 2013). For these reasons, I would argue that play behaviour has positive welfare outcomes. Whether the horses in my study demonstrated positive welfare after play still requires further testing of the carry over effects of play over a longer time period.

In future, it would be important to consider the timing of salivary cortisol samples in a variety of species when they are exposed to a stressor. Then one could assess whether salivary cortisol has an underlying physiological role with the positive behaviours observed in instances where animals display playful behaviours. It would also be useful to incorporate other physiological measures like faecal cortisol or heart rate, in conjunction with play behaviour and decision-making, to determine if there is any form of physiological and behavioural link in how animals respond.

REFERENCES

Anderson, C., Keyserlingk, M. von, Lidfors, L.M., Weary, D.M., 2020. Anticipatory behaviour in animals: A critical review. *Animal Welfare* 29, 231–238. https://doi.org/10.7120/09627286.29.3.231

Andrade, M.M.M., Tomé, M.F., Santiago, E.S., Lúcia-Santos, A., de Andrade, T.G.C.S., 2003. Longitudinal study of daily variation of rats' behaviour in the elevated plus-maze. *Physiology & Behaviour* 78, 125–133. https://doi.org/10.1016/S0031-9384(02)00941-1

Antonevich, A.L., Alekseeva, G.S., Vasilieva, N.A., Pavlova, E.V., Loshchagina, J.A., Duplyakina, S.Yu., Naidenko, S.V., 2019. Social play changes reflect differences in biology and development of three felids. *Russian Journal Theriology* 18, 80–90. https://doi.org/10.15298/rusjtheriol.18.2.02

Apfelbach, R., Blanchard, C.D., Blanchard, R.J., Hayes, R.A., McGregor, I.S., 2005. The effects of predator odours in mammalian prey species: A review of field and laboratory studies. *Neuroscience & Biobehavioural Reviews, Defensive Behaviour* 29, 1123–1144. https://doi.org/10.1016/j.neubiorev.2005.05.005

Baragli, P., Vitale, V., Banti, L., Sighieri, C., 2014. Effect of aging on behavioural and physiological responses to a stressful stimulus in horses (*Equus caballus*). *Behaviour* 151, 1513–1533. https://doi.org/10.1163/1568539X-00003197

Barrett, H.C., 2015. Adaptations to predators and prey, in: The handbook of evolutionary psychology. *John Wiley & Sons, Ltd*, 200–223. https://doi.org/10.1002/9780470939376.ch7

Bartolomé, E., Cockram, M.S., 2016. Potential effects of stress on the performance of sport horses. *Journal of Equine Veterinary Science* 40, 84–93. https://doi.org/10.1016/j.jevs.2016.01.016

Bekoff, M., 1984. Social play behaviour. *Bioscience*, 34(4), 228-233.

Boonstra, R., 2004. Coping with changing northern environments: The role of the stress axis in birds and mammals. *Integrative and Comparative Biology* 44, 95–108. https://doi.org/10.1093/icb/44.2.95

Bowman, R.E., Beck, K.D., Luine, V.N., 2003. Chronic stress effects on memory: sex differences in performance and monoaminergic activity. *Hormones and Behaviour* 43, 48–59. https://doi.org/10.1016/S0018-506X(02)00022-3

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

Burgdorf, J., Kroes, K.A., Moskal, J.R., 2017. Rough and tumble play induces resilience to stress in rats. *NeuroReport* 28 (17), 1122 – 1126. https://oce.ovid.com/article/00001756-201712020-00003/PDF

Burghardt, G.M., 2005. *The Genesis of Animal Play: Testing the Limits.* MIT Press.

Burghardt, G.M., Dinets, V., Murphy, J.B., 2015. Highly repetitive object play in a cichlid fish (*Tropheus duboisi*). *Ethology* 121, 38–44. https://doi.org/10.1111/eth.12312

Christensen, J.W., Zharkikh, T., Chovaux, E., 2011. Object recognition and generalisation during habituation in horses. *Applied Animal Behaviour Science* 129, 83–91. https://doi.org/10.1016/j.applanim.2010.11.016

Christiansen, J.J., Djurhuus, C.B., Gravholt, C.H., Iversen, P., Christiansen, J.S., Schmitz, O., Weeke, J., Jørgensen, J.O.L., Møller, N., 2007. Effects of cortisol on carbohydrate, lipid, and protein metabolism: Studies of acute cortisol withdrawal in adrenocortical failure. *The Journal of Clinical Endocrinology & Metabolism* 92, 3553–3559. https://doi.org/10.1210/jc.2007-0445

Clark, C. D., Miller, P. J., 1998. Play. In H. Friedman (ed.), *Encyclopedia of Mental Health* 3, 189–197. San Diego: Academic Press.

Clark, F., Smith, L., 2013. Effect of a cognitive challenge device containing food and non-food rewards on chimpanzee well-being. *American journal of primatology* 75. https://doi.org/10.1002/ajp.22141

Dai, F., Costa, A., Bonfanti, L., Caucci, C., Di Martino, G., Lucarelli, R., Padalino, B., Minero, M., 2019. Positive reinforcement-based training for self-loading of meat horses reduces loading time and stress-related behaviour. *Frontiers in Veterinary Science* 6. https://doi.org/10.3389/fvets.2019.00350

DeVries, A.C., Glasper, E.R., Detillion, C.E., 2003. Social modulation of stress responses. *Physiology & Behaviour, A Tribute to Paul MacLean: The Neurobiological Relevance of Social Behaviour* 79, 399–407. https://doi.org/10.1016/S0031-9384(03)00152-5

Dias-Ferreira, E., Sousa, J.C., Melo, I., Morgado, P., Mesquita, A.R., Cerqueira, J.J., Costa, R.M., Sousa, N., 2009. Chronic stress causes frontostriatal reorganization and affects decision-making. *Science* 325, 621–625. https://doi.org/10.1126/science.1171203

Donaldson, T.M., Newberry, R.C., Špinka, M., Cloutier, S., 2002. Effects of early play experience on play behaviour of piglets after weaning. *Applied Animal Behaviour Science* 79, 221–231. https://doi.org/10.1016/S0168-1591(02)00138-7

Donelan, S.C., Trussell, G.C., 2020. Sex-specific differences in the response of prey to predation risk. *Functional Ecology* 34, 1235–1243. https://doi.org/10.1111/1365-2435.13569

Ellis, A.D., Stephenson, M., Preece, M., Harris, P., 2014. A novel approach to systematically compare behavioural patterns between and within groups of horses. *Applied Animal Behaviour Science* 161, 60–74. https://doi.org/10.1016/j.applanim.2014.09.017

Esch, L., Wöhr, C., Erhard, M., Krüger, K., 2019. Horses' (*Equus Caballus*) laterality, stress hormones, and task related behaviour in innovative problem-solving. *Animals (Basel)* 9, 265. https://doi.org/10.3390/ani9050265

Grace, L., Hescham, S., Kellaway, L.A., Bugarith, K., Russell, V.A., 2009. Effect of exercise on learning and memory in a rat model of developmental stress. *Metabolic Brain Disease* 24, 643–657. https://doi.org/10.1007/s11011-009-9162-5

Guimarães-Costa, R., Guimarães-Costa, M.B., Pippa-Gadioli, L., Weltson, A., Ubiali, W.A., Paschoalin-Maurin, T., Felippotti, T.T., Elias-Filho, D.H., Laure, C.J., Coimbra, N.C., 2007. Innate defensive behaviour and panic-like reactions evoked by rodents during aggressive encounters with Brazilian constrictor snakes in a complex labyrinth: Behavioural validation of a new model to study affective and agonistic reactions in a prey versus predator paradigm. *Journal of Neuroscience Methods* 165, 25–37. https://doi.org/10.1016/j.jneumeth.2007.05.023

Hall, C., Randle, H., Pearson, G., Preshaw, L., Waran, N., 2018. Assessing equine emotional state. *Applied Animal Behaviour Science* 205, 183–193. https://doi.org/10.1016/j.applanim.2018.03.006

Hall, K., Bryant, J., Staley, M., Whitham, J.C., Miller, L.J., 2021. Behavioural diversity as a potential welfare indicator for professionally managed chimpanzees (*Pan troglodytes*): Exploring variations in calculating diversity using species-specific behaviours. *Animal Welfare* 30, 381–392. https://doi.org/10.7120/09627286.30.4.002

Han, C.S., Jablonski, P.G., Brooks, R.C., 2015. Intimidating courtship and sex differences in predation risk lead to sex-specific behavioural syndromes. *Animal Behaviour* 109, 177–185. https://doi.org/10.1016/j.anbehav.2015.08.017

Hausberger, M., Fureix, C., Bourjade, M., Wessel-Robert, S., Richard-Yris, M.-A., 2012. On the significance of adult play: what does social play tell us about adult horse welfare? *Naturwissenschaften* 99, 291–302. https://doi.org/10.1007/s00114-012-0902-8

Held, S.D.E., Špinka, M., 2011. Animal play and animal welfare. *Animal Behaviour* 81, 891–899. https://doi.org/10.1016/j.anbehav.2011.01.007

Holtby, A.R., McGivney, B.A., Browne, J.A., Katz, L.M., Murphy, K.J., Hill, E.W., 2023. Variation in salivary cortisol responses in yearling Thoroughbred racehorses during their first year of training. *PLOS One* 18, e0284102. https://doi.org/10.1371/journal.pone.0284102

Kang, O.-D., Lee, W.-S., 2016. Changes in salivary cortisol concentration in horses during different types of exercise. *Asian-Australasian Journal of Animal Science* 29, 747–752. https://doi.org/10.5713/ajas.16.0009

Koolhaas, J.M., Bartolomucci, A., Buwalda, B., de Boer, S.F., Flügge, G., Korte, S.M., Meerlo, P., Murison, R., Olivier, B., Palanza, P., Richter-Levin, G., Sgoifo, A., Steimer, T., Stiedl, O., van Dijk, G., Wöhr, M., Fuchs, E., 2011. Stress revisited: a critical evaluation of the stress concept. *Neuroscience and Biobehavioural Reviews* 35, 1291–1301. https://doi.org/10.1016/j.neubiorev.2011.02.003

Krüger, K., Heinze, J., 2008. Horse sense: Social status of horses (*Equus caballus*) affects their likelihood of copying other horses' behaviour. *Animal cognition* 11, 431–9. https://doi.org/10.1007/s10071-007-0133-0

Lagos, P.A., Herberstein, M.E., 2017. Are males more scared of predators? Differential change in metabolic rate between males and females under predation risk. *Physiology & Behaviour* 173, 110–115. https://doi.org/10.1016/j.physbeh.2017.02.002

Lapuschkin, S., Wäldchen, S., Binder, A., Montavon, G., Samek, W., Müller, K.-R., 2019. Unmasking Clever Hans predictors and assessing what machines really learn. *Nature Communications* 10, 1096. https://doi.org/10.1038/s41467-019-08987-4

Leiner, L., Fendt, M., 2011. Behavioural fear and heart rate responses of horses after exposure to novel objects: Effects of habituation. *Applied Animal Behaviour Science* 131, 104–109. https://doi.org/10.1016/j.applanim.2011.02.004

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. https://doi.org/10.1371/journal.pone.0010257

Lima, A.F. de, Lima, S.G.C., Nogueira-Filho, S.L.G., Held, S.D.E., Mendl, M., Nogueira, S.S.C., 2024. Object play as a positive emotional state indicator for farmed spotted paca (*Cuniculus paca*). *Animals* 14, 78. https://doi.org/10.3390/ani14010078

Lonsdorf, E.V., Markham, A.C., Heintz, M.R., Anderson, K.E., Ciuk, D.J., Goodall, J., Murray, C.M., 2014. Sex differences in wild chimpanzee behaviour emerge during infancy. *PLOS One* 9, e99099. https://doi.org/10.1371/journal.pone.0099099

Loy, J., 1970. Behavioural responses of free-ranging rhesus monkeys to food shortage. *American Journal of Physical Anthropology* 33, 263–271. https://doi.org/10.1002/ajpa.1330330212

Lundberg, P., Hartmann, E., Roth, L.S.V., 2020. Does training style affect the human-horse relationship? Asking the horse in a separation–reunion experiment with the owner and a stranger. *Applied Animal Behaviour Science* 233, 105144. https://doi.org/10.1016/j.applanim.2020.105144

McDonnell, S.M., Poulin, A., 2002. Equid play ethogram. *Applied Animal Behaviour Science, Equine Behaviour* 78, 263–290. https://doi.org/10.1016/S0168-1591(02)00112-0

McEwen, B.S., Karatsoreos, I.N., 2020. *What Is Stress?* Springer International Publishing, 19–42. https://doi.org/10.1007/978-3-030-16996-1_4

McEwen, B.S., Sapolsky, R.M., 1995. Stress and cognitive function. *Current Opinion in Neurobiology* 5, 205–216. https://doi.org/10.1016/0959-4388(95)80028-X

Mendl, M., 1999. Performing under pressure: stress and cognitive function. *Applied Animal Behaviour Science* 65, 221–244. https://doi.org/10.1016/S0168-1591(99)00088-X

Metz, G.A.S., Schwab, M.E., Welzl, H., 2001. The effects of acute and chronic stress on motor and sensory performance in male Lewis rats. *Physiology & Behaviour* 72, 29–35. https://doi.org/10.1016/S0031-9384(00)00371-1

Mika, A., Mazur, G.J., Hoffman, A.N., Talboom, J.S., Bimonte-Nelson, H.A., Sanabria, F., Conrad, C.D., 2012. Chronic stress impairs prefrontal cortex-dependent response inhibition and spatial working memory. *Behavioural Neuroscience* 126, 605–619. https://doi.org/10.1037/a0029642

Miller, L.J., Lauderdale, L.K., Bryant, J.L., Mellen, J.D., Walsh, M.T., Granger, D.A., 2021. Behavioural diversity as a potential positive indicator of animal welfare in bottlenose dolphins. *PLOS One* 16, e0253113. https://doi.org/10.1371/journal.pone.0253113

Miller, L.J., Vicino, G.A., Sheftel, J., Lauderdale, L.K., 2020. Behavioural diversity as a potential indicator of positive animal welfare. *Animals* 10, 1211. https://doi.org/10.3390/ani10071211

Morey, J.N., Boggero, I.A., Scott, A.B., Segerstrom, S.C., 2015. Current directions in stress and human immune function. *Current Opinion in Psychology* 5, 13–17. https://doi.org/10.1016/j.copsyc.2015.03.007

Möstl, E., Palme, R., 2002. Hormones as indicators of stress. *Domestic Animal Endocrinology* 23, 67–74. https://doi.org/10.1016/s0739-7240(02)00146-7

Mowrey, W.R., Portman, D.S., 2012. Sex differences in behavioral decision-making and the modulation of shared neural circuits. *Biology of Sex Differences* 3, 8. https://doi.org/10.1186/2042-6410-3-8

Murphy, E., Nordquist, R.E., van der Staay, F.J., 2014. A review of behavioural methods to study emotion and mood in pigs, *Sus scrofa*. *Applied Animal Behaviour Science* 159, 9–28. https://doi.org/10.1016/j.applanim.2014.08.002

O'Connor, E.A., Parker, M.O., McLeman, M.A., Demmers, T.G., Lowe, J.C., Cui, L., Davey, E.L., Owen, R.C., Wathes, C.M., Abeyesinghe, S.M., 2010. The impact of chronic environmental stressors on growing pigs, *Sus scrofa* (Part 1): Stress physiology, production and play behaviour. *Animal* 4, 1899–1909. https://doi.org/10.1017/S1751731110001072

Olczak, K., Winther Christensen, J., Klocek, C., 2018. Food motivation in horses appears stable across different test situations. *Applied Animal Behaviour Science* 204, 60–65. https://doi.org/10.1016/j.applanim.2018.04.006

Oliveira, A.F.S., Rossi, A.O., Silva, L.F.R., Lau, M.C., Barreto, R.E., 2010. Play behaviour in nonhuman animals and the animal welfare issue. *Journal of Ethology* 28, 1–5. https://doi.org/10.1007/s10164-009-0167-7

Pellis, S., Pellis, V., 2013. *The Playful Brain: Venturing to the Limits of Neuroscience*. Simon and Schuster.

Pellis, S.M., Pellis, V.C., 2017. What is play fighting and what is it good for? *Learning and Behaviour* 45, 355–366. https://doi.org/10.3758/s13420-017-0264-3

Piyapong, C., Krause, J., Chapman, B.B., Ramnarine, I.W., Louca, V., Croft, D.P., 2010. Sex matters: a social context to boldness in guppies (*Poecilia reticulata*). *Behavioural Ecology* 21, 3–8. https://doi.org/10.1093/beheco/arp142

Pyke, G.H., Starr, C.K., 2021. *Optimal Foraging Theory, Encyclopaedia of Social Insects.* Springer International Publishing 677–685. https://doi.org/10.1007/978-3-030-28102-1_90

Rabin, L.A., 2003. Maintaining behavioural diversity in captivity for conservation: Natural behaviour management. *Animal Welfare* 12, 85–94. https://doi.org/10.1017/S0962728600025409

Rama, E., Capi, G., Mizusaki, S., Tanaka, N., Kawahara, S., 2016. Effects of environment changes on rat`s learned behaviour in an elevated Y-maze. *Journal of Medical and Bioengineering* 5, 113–118. https://doi.org/10.18178/jomb.5.2.113-118

Réale, D., Reader, S.M., Sol, D., McDougall, P.T., Dingemanse, N.J., 2007. Integrating animal temperament within ecology and evolution. *Biological Reviews* 82, 291–318. https://doi.org/10.1111/j.1469-185X.2007.00010.x

Safryghin, A., Hebesberger, D.V., Wascher, C.A.F., 2019. Testing for behavioural and physiological responses of domestic horses (*Equus caballus*) across different contexts – consistency over time and effects of context. *Frontiers in Psychology* 10.

Samhita, L., Gross, H.J., 2013. The "Clever Hans Phenomenon" revisited. *Communicative & Integrative Biology* 6, e27122. https://doi.org/10.4161/cib.27122

Sandi, C., 2013. Stress and cognition. *WIREs Cognitive Science* 4, 245–261. https://doi.org/10.1002/wcs.1222

Sapolsky, R.M., 2002. *Endocrinology of the stress-response*. Behavioral Endocrinology, 2nd Ed. MIT Press, Cambridge, MA, US, 409–450.

Sauer, F.J, Gerber, V., Frei, S., Bruckmaier, R.M., Groessl, M., 2020. Salivary cortisol measurement in horses: immunoassay or LC-MS/MS? *PubMed*. https://pubmed.ncbi.nlm.nih.gov/32247992/

Schubiger, M.N., Fichtel, C., Burkart, J.M., 2020. Validity of cognitive tests for non-human animals: Pitfalls and prospects. *Frontiers in Psychology* 11.

Selye, H., 1936. A syndrome produced by diverse nocuous agents. *Nature* 138, 32–32. https://doi.org/10.1038/138032a0

Sherwin, C.M., 2004. The influences of standard laboratory cages on rodents and the validity of research data. *Animal Welfare* 13, S9–S15. https://doi.org/10.1017/S0962728600014329

Shettleworth, S.J., 2001. Animal cognition and animal behaviour. *Animal Behaviour* 61, 277–286. https://doi.org/10.1006/anbe.2000.1606

Sigurdsson, T., Duvarci, S., 2016. Hippocampal-prefrontal interactions in cognition, behaviour and psychiatric disease. *Frontiers in Systems Neuroscience* 9.

Soravia, C., Ashton, B.J., Thornton, A., Ridley, A.R., 2021. The impacts of heat stress on animal cognition: Implications for adaptation to a changing climate. *WIREs Climate Change* 12, e713. https://doi.org/10.1002/wcc.713

Starcke, K., Brand, M., 2012. Decision making under stress: A selective review. *Neuroscience & Biobehavioural Reviews* 36, 1228–1248. https://doi.org/10.1016/j.neubiorev.2012.02.003

Tao of Horsemanship Academy, 2020. *Holistic Horsemanship*. https://www.taoofhorsemanship.com

Thomas, P., Robertson, L., 1991. Plasma cortisol and glucose stress responses of red drum (*Sciaenops ocellatus*) to handling and shallow water stressors and anaesthesia with MS-222, quinaldine sulphate and etomidate. *Aquaculture* 96, 69–86. https://doi.org/10.1016/0044-8486(91)90140-3

Tinbergen, N., 1963. On aims and methods of Ethology. *Zeitschrift für Tierpsychologie* 20, 410–433. https://doi.org/10.1111/j.1439-0310.1963.tb01161.x

Torcivia, C., McDonnell, S., 2021. Equine discomfort ethogram. *Animals (Basel)* 11, 580. https://doi.org/10.3390/ani11020580

Vanderschuren, L.J., Niesink, R.J., Van Ree, J.M., 1997. The neurobiology of social play behaviour in rats. *Neuroscience and Biobehavioural Reviews* 21, 309–326. https://doi.org/10.1016/s0149-7634(96)00020-6

Verbeek, E., Colditz, I., Blache, D., Lee, C., 2019. Chronic stress influences attentional and judgement bias and the activity of the HPA axis in sheep. *PLOS One* 14, e0211363. https://doi.org/10.1371/journal.pone.0211363

Verbruggen, F., Logan, G.D., 2008. Response inhibition in the stop-signal paradigm. *Trends in Cognitive Sciences* 12, 418–424. https://doi.org/10.1016/j.tics.2008.07.005

Wallace, K.J., Rosen, J.B., 2000. Predator odour as an unconditioned fear stimulus in rats: Elicitation of freezing by trimethylthiazoline, a component of fox faeces. *Behavioural Neuroscience* 114, 912–922. https://doi.org/10.1037/0735-7044.114.5.912

Wang, S., Aamodt, S., 2012. Play, stress, and the learning brain. *Cerebrum* 12.

Waring, G,. 2003. *Horse behaviour*, 2nd edition. Noyes Publications. William Andrew Publishing, Norwich.

Wechsler, B., 1995. Coping and coping strategies: a behavioural view. *Applied Animal Behaviour Science* 43, 123–134. https://doi.org/10.1016/0168-1591(95)00557-9

Wemelsfelder, F., Haskell, M., Mendl, M.T., Calvert, S., Lawrence, A.B., 2000. Diversity of behaviour during novel object tests is reduced in pigs housed in substrate-impoverished conditions. *Animal Behaviour* 60, 385–394. https://doi.org/10.1006/anbe.2000.1466

Whirledge, S., Cidlowski, J.A., 2010. Glucocorticoids, stress, and fertility. *Minerva Endocrinology* 35, 109–125.

Wilson, S.C., Kleinman, D.G., 1974. Eliciting play: A comparative study. *American Zoologist* 14, 341–370. https://doi.org/10.1093/icb/14.1.341

Yamanashi, Y., Nogami, E., Teramoto, M., Morimura, N., Hirata, S., 2018. Adult-adult social play in captive chimpanzees: Is it indicative of positive animal welfare? *Applied Animal Behaviour Science* 199, 75–83. https://doi.org/10.1016/j.applanim.2017.10.006

Yarnell, K., Hall, C., Royle, C., Walker, S.L., 2015. Domesticated horses differ in their behavioural and physiological responses to isolated and group housing. *Physiology & Behaviour* 143, 51–57. https://doi.org/10.1016/j.physbeh.2015.02.040

Yerkes, R.M., Dodson, J.D., 1908. The relation of strength of stimulus to rapidity of habit-formation. *Journal of Comparative Neurology and Psychology* 18, 459–482. https://doi.org/10.1002/cne.920180503