

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

Jodi Anne Howard^{*} , Neville Pillay

School of Animal, Plant, and Environmental Sciences, University of the Witwatersrand, Private Bag 3, WITS 2050, South Africa

ARTICLE INFO

Keywords:

Cognition
Decision-making
Horses
Object play behaviour
Stress
Stressor

ABSTRACT

Horses (*Equus caballus*) are prone to acute and chronic stress, particularly in equestrian sports, raising significant welfare concerns. We investigated whether and how play can mitigate acute acoustic stress and enhance decision-making in horses. We studied horses in two different stabling and working conditions (intensively vs. extensively managed) and recorded their behaviours in a Y maze for a food reward in three sequential treatments: baseline (no stressor), stress (exposure to a short-term acoustic stressor) and solitary object play with a large ball (exposure to the stressor followed by play). We recorded behavioural diversity, latency to approach the decision point (before choice) and food reward arm, behaviours associated with positive and negative emotional state and overall choices made (decisions) in the maze. Decision-making was significantly better in the play treatment than in both the baseline and stress treatments. Horses in the stress treatment exhibited a greater diversity of behaviours associated with negative emotional states and poorer decision-making compared to the play and baseline treatments. Additionally, intensively managed horses under more constrained conditions, were more stressed than those extensively managed. Our findings suggest that play can effectively reduce stress and improve decision-making in horses, depending on their housing conditions, and highlight the importance of play for equine welfare.

1. Introduction

Stress is the suite of behavioural and physiological responses an animal mounts when a stressor (i.e. any internal or external challenge) threatens (or is perceived to threaten) homeostasis (Selye, 1936; Sapolsky, 2002). When the stressor is judged to be dangerous, the sympathetic-adreno-medullary 'fight-or-flight' cascade releases catecholamines and glucocorticoids such as cortisol (Möstl and Palme, 2002; Bartolomé and Cockram, 2016). The immediate affective output is an emotional response (i.e. a transient state defined by valence and arousal); across repeated episodes, these moments accumulate into a longer-term emotional state (Mendl et al., 2010). Briefly, predictable stressors can induce eustress (positive stress) and focus attention (Yerkes and Dodson, 1908; Lupien et al., 2007) and, in rodents, enhanced hippocampal plasticity and remote memory (Kirby et al., 2013). For example, before a race, thoroughbred racehorses experience short-term or eustress due to the release of stress hormones and can show enhanced performance (reviewed in Bartolomé and Cockram, 2016). Prolonged or unpredictable HPA activation, by contrast, promotes

immunosuppression, stereotypy and hippocampal dysfunction (McEwen and Sapolsky, 1995; Möstl and Palme, 2002). Because glucocorticoids modulate synaptic plasticity in hippocampal-pre-frontal circuits, decision making quality can shift within minutes of stress onset (Dias-Ferreira et al., 2009). Buffering such transient cognitive disruptions is critical for equine welfare in modern yards where sudden stressors, such as loud machinery, are common.

Cognition entails integrating environmental information and selecting appropriate actions (Shettleworth, 2001; Sandi, 2013). In mammals, the hippocampal-pre-frontal loops that support spatial memory and behavioural flexibility are especially sensitive to stress hormones (Sigurdsson and Duvarci, 2016). Persistent glucocorticoid elevation therefore impairs learning and flexible choice (Bowman et al., 2003; Mika et al., 2012), whereas performance peaks under mild arousal (Yerkes and Dodson, 1908). Management strategies must thus down-regulate distress without dampening persistent alertness.

Attempts to relieve stress in captive or domesticated animals typically focus either on removing the stressor (e.g. improving housing stability) or on providing environmental enrichment that boosts

^{*} Corresponding author.

E-mail address: jodimol1330@gmail.com (J.A. Howard).

<https://doi.org/10.1016/j.applanim.2025.106752>

Received 5 March 2025; Received in revised form 10 July 2025; Accepted 14 July 2025

Available online 15 July 2025

0168-1591/© 2025 The Author(s). Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

physical activity (e.g. turnout, exercise) and mental stimulation (e.g. puzzle feeders – for horses, slow-release hay nets and salt-lick blocks are widely used examples), thereby reducing boredom and enhancing cognition (Coleman and Novak, 2017). Recent work shows that enrichment can also promote positive affect: short, daily interventions increased play and ‘relaxed’ postures in riding-school horses (Bini de Lima et al., 2023). Such interventions can be implemented in the short-term to deflect acute stressors or sustained in the long-term to mitigate chronic stress (Shepherdson, 1994).

Play, particularly object play, is a rapid, handler-independent enrichment that can reverse acute stress effects. For example, adult horses manipulate balls or barrels more when natural foraging and social outlets are restricted (Hausberger et al., 2012; Christensen et al., 2011); play frequency is inversely correlated with hair-cortisol (Mazzola et al., 2021), and in several species, play recruits parasympathetic pathways linked to affective recovery (Held and Špinka, 2011). Across mammals and birds, brief play bouts have been shown to reduce behavioural and neuro-endocrine stress indices (rats: van den Berg et al., 1999; chickens: Lundén et al., 2022), while in children play therapy accelerates cortisol recovery after acute stress (Potas et al., 2013). For ungulates housed in barren boxes, object play may substitute for grazing, sparring and locomotion, providing a behavioural release valve that restores affective balance (Meehan and Mench, 2007; Hausberger et al., 2012). Play is also linked to enhanced synaptic plasticity and faster reversal learning in several mammals, suggesting a plausible cognitive ‘reset’ mechanism that could translate into more accurate decision-making (Pellis and Pellis, 2013; Held and Špinka, 2011).

Horses are often subjected to stress because of routine practices in the equestrian industry. Riding-school horses, for instance, may experience chronic stress from frequent riding by multiple riders, leading to compromised welfare and an increase in play behaviour as a form of escape (Hausberger et al., 2012). Likewise, horses kept for leisure or for sports such as dressage, show-jumping and racing can face significant stress owing to their demanding roles (Bartolomé and Cockram, 2016; Becker-Birck et al., 2013; Esch et al., 2019). Yet, no study has tested whether a brief bout of object play can modify decision-making in horses after an acute anthropogenic stressor.

We aimed to establish whether object play could improve decision-making accuracy and emotional state following a short acoustic stressor. Horses from two contrasting management systems (i.e. an intensively managed riding school and an extensively managed pasture operation) were exposed to a 30-s acoustic stressor and then tested in a Y-maze offering a food versus a non-food choice. Horses can discriminate quantities and negotiate detour problems, demonstrating complex cognitive abilities (Hanggi, 2003; Savin and Randle, 2013; Valençon et al., 2013). In a second treatment, each horse interacted with a familiar 1-m diameter play ball for two minutes immediately after the stressor before being retested in the maze. We predicted that (i) object play would improve Y-maze decision accuracy and produce a more positive emotional state relative to the stressor-only condition, and (ii) decision accuracy would be lower and emotion more negative in pasture-managed horses because these animals encounter anthropogenic noise less frequently and therefore show stronger startle and cortisol responses (Christensen et al., 2006) than in their riding-school counterparts.

2. Materials and method

2.1. Study sites

Horses at two sites were selected for study, located 54 km apart in the Gauteng Province, South Africa. The first site, denoted as Centre A, was a private farm with few houses in the vicinity. The second site, was a horse farm open to the public and was surrounded by houses, denoted as Centre B. Centre A had extensive, pasture-based management, whereas Centre B adopted an intensive, riding-school management. We obtained

ethics approval for our study from the University of the Witwatersrand Animal Research and Ethics Committee; clearance number AREC23-03-018B.

2.2. Study subjects

Twenty horses were selected for study, ten from Centre A and ten from Centre B. Three horses died during the study: two from Centre A due to African Horse Sickness and one from Centre B which was euthanised following a fractured leg that became septic. The data for these horses were excluded from the final analyses. All remaining horses completed the full experimental protocol. Centre A (Extensive) horses averaged 10.5 ± 4.9 years (range 5–16) and were all Boer ponies (6 geldings, 2 mares) working an estimated 1–2 h per week for fence-line patrol and cattle herding. They lived in a 24-h mixed herd on pasture, with grazing *ad libitum* and a maize supplement every few days in winter. They were unshod, their hooves were trimmed twice a year, and veterinary care was on call. Formal back-health screening was not undertaken, but no back-pain was noted during handling or testing.

Centre B (Intensive) horses averaged 14.4 ± 5.4 years (range 7–24) and comprised of five Thoroughbreds, three cross-breeds and one Clydesdale (6 geldings, 3 mares). They worked an estimated 8–10 h per week in lessons, were housed overnight in 3×3 m box stalls bedded with wood shavings (changed every second day) and turned out to pasture for 6 h daily. Feeding consisted of three concentrate meals plus hay nets refreshed twice daily and pasture grazing during turnout. Social contact was provided via adjacent stalls and partial group turnout. Veterinary checks were done monthly; hooves were shod and reset every six weeks. No back-pain was detected during the pre-study check. The horses would normally be removed from work at any sign of back discomfort, which was not detected in this study.

We conducted *ad libitum* observations of the horses in their paddocks before the start of the experiments at each location to construct an ethogram for behaviours.

2.3. Food choice test

Prior to experiments, each horse underwent a food choice test. Each horse was individually presented with one cup (250 ml) each of three food types, carrots, apples and horse lucerne pellets (considered treats for horses) in separate plastic feeding bowls (25 cm by 15 cm by 10 cm). The feeding bowls comprised of different colours randomly assigned for each food test and were placed one metre apart. Each horse was allowed 10 min to sample the three food types. The food type that was consumed and finished first was regarded as the most preferred food type that for each horse, which was then used as the incentive during the decision-making tests later.

2.4. 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 (1.5 m in height) and three lengths of rope (2 of 6 m in length, 2 of 3.7 m in length and 2 of 3 m in length) (Fig. 1). The wooden poles were sunk into the ground and the ropes were secured to the tops of the wooden poles with metal rings. A 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 fifteen minutes per horse for two weeks before the start of the experiments to acclimate the horses to the maze (Christensen et al., 2011; Leiner and Fendt, 2011). Each horse was also individually exposed to the 1 m diameter ball that was later used in the play treatment. Horses were exposed to the ball during four 15-min sessions per day over a two-week period in a round pen (16 m in diameter) to ensure familiarity and to prevent novelty responses during the actual play trials. All horses interacted with the ball

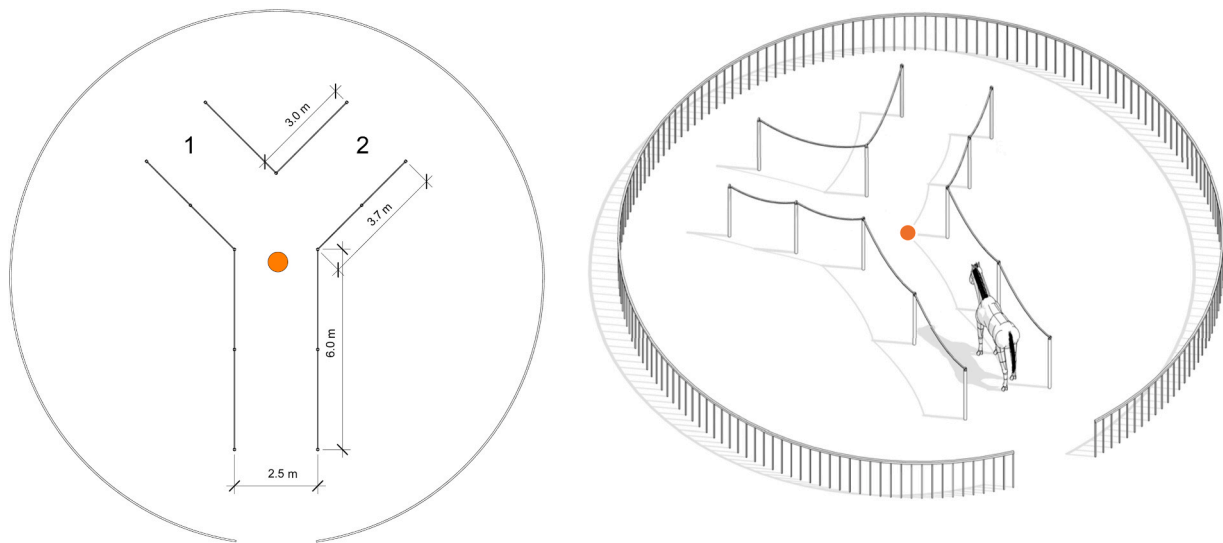


Fig. 1. Schematic diagrams of the Y-maze within a round pen used in tests. The horse entered the maze at the bottom and moved toward the orange dot. The orange dot represents the decision point where the horse had to decide to move to exit 1 or 2.

freely during these sessions without prompting.

2.5. Treatments

Each horse underwent three treatments, namely baseline, stress and play. Each horse was tested twice in the treatments, with a minimum of one week between repeat experimental treatments (Supplementary Figure 1). 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. The horse's behaviours were recorded using a smartphone (Samsung Galaxy A50) in each treatment.

Each horse completed two experimental blocks, separated by at least seven days; each block comprised two test days, Day 1 (Baseline followed by Stress) and Day 2 (Baseline followed by Play). This fixed sequence, Stress first, then Play, was essential because the study aimed to ascertain whether object play could buffer a previously induced stress response; reversing the order would have confounded that question. Baseline trials were always run first on each test day to obtain a same-day reference, and a minimum interval of 48 h (twice the 24-h baseline recovery in behaviour) separated the Stress and Play days, minimising carry-over effects (timeline illustrated in Supplementary Fig. 1). All sessions were video-recorded with a Samsung Galaxy A50 positioned 4 m from the maze. To check for any week-to-week sequence artefact, decision accuracy was re-analysed in a linear mixed model with "Block" (first vs. second) as a fixed factor and horse ID as a random effect; Block had no effect ($F_{1,16} = 0.42$, $p = 0.530$), confirming that the 48-h buffer and between-block interval were sufficient to eliminate sequence bias.

2.6. Baseline treatment

Each horse was placed in a round pen (16 m in diameter) with the Y-maze, in which one arm had each horse's preferred food type and the other arm contained the rocks. The position of the food bowl vs. the rock bowl was randomly assigned for each trial. The test arena was large and open, making it unlikely that horses could simply detect the reward by olfactory cues from a distance without actively investigating the bowls. Nonetheless, our study focused on whether horses' decision-making following an acute stressor or play event would alter their approach to the decision task, in a familiar context where the reward and bowl cues were already known. The rocks were used in the feeding bowl of one of

the arms as a neutral scent and not an incentive.

Using an electronic stopwatch we measured initial latency (maze entry to the decision point) and final latency (decision point to reaching a bowl). Decisions were scored "yes" when the horse chose the arm with its preferred food and "no" when it chose the rock-filled arm. From the video recordings, JAH recorded every posture change continuously and later collapsed the data into 10-s bins, recording eight mutually-exclusive postures for analysis (Table 2) and, for completeness, continuously noted overt behaviours (snort, paw, buck, rear, kick, freeze), although these were not included in the valence metric. We recorded postures, rather than grouping behaviours into functional categories, to avoid biasing the classification of emotional states.

Each horse was observed for three 5 min periods (pre-maze, in-maze, post-maze), yielding 30 bins per period. Behavioural diversity was calculated as the count of distinct posture states in the bins before the decision point and in the bins from the decision point to the end of the maze. Emotional valence followed the ear-eye-neck coding of Torcivia and McDonnell (2021) and Waring (2003). With 30 scans per 5-min period, a summed score of +10 indicated that the horse displayed positive ear-eye-neck cues in ≥ 20 scans ($\geq 67\%$), whereas a score of -10 indicated the same proportion of negative cues. This two-thirds-majority threshold strikes a balance between sensitivity and reliability—high enough to minimise misclassification due to momentary postural shifts, yet not so strict as to obscure genuine affective changes. Emotional state was assigned for the full maze travel. A decision was deemed successful when the horse chose the preferred-food arm at both the initial head-position decision point (Fig. 1) and the final bowl.

2.7. Stress treatment

Five minutes after the baseline test, two 5-s anthropogenic sounds were broadcast consecutively, motorbike followed by drone, for a total exposure of 10-s. The motorbike stimulus was a 125 cc two-stroke trail bike accelerating, recorded at a distance of 3 m with a Zoom H4n recorder (44.1 kHz, 16-bit; dominant frequency band 120–450 Hz). The drone stimulus was a DJI Mavic Air hovering, recorded 30 cm beneath the rotors with the same recorder (dominant band 2–3 kHz). Both clips were replayed through a Volcano Bluetooth speaker positioned 1 m outside the round pen and facing inward; peak A-weighted sound-pressure levels measured at horse-ear height (1.6 m) were 83 ± 2 dB for the motorbike and 76 ± 2 dB for the drone. The paired stimuli mirrored common show-ground noise and provide complementary low-

and mid-frequency energy. Pilot trials confirmed salience: horses showed a startle response (ear-flick, head-raise) without panic. Immediately after exposure, each horse entered the decision-making maze, and latency, posture diversity, emotional valence and initial and final choices were recorded as described above.

2.8. Play treatment

After the baseline treatment in the round pen, the recorded motor-bike and drone sounds were played from a smartphone as described above. The ball (described above) was then placed in the round pen. The horse was allowed to play with the ball, unaided, for 5 min, and its behaviour was recorded. Thereafter, the horse underwent the decision-making test, as mentioned above. Each horse’s behaviours were recorded during the experiment and then continuously after the experiment for 5 min (Table 1). The latency, behavioural frequency, emotional states, initial decision and final decision were calculated, as described above.

Immediately after the sound exposure, the familiar ball (see Habituation to the Experimental Set-up) was placed in the round pen. Each horse was then allowed to play with the ball, unaided, for 5 min, and its behaviour was recorded. Horses typically engaged with the ball by nosing, pushing and pawing, consistent with object-directed play. The horses were consistently assessed as being relaxed based on our pre-defined emotional state criteria, and no stress-related behaviours (e.g. galloping, tail-flagging, snorting) were observed. Each horse’s behaviours were recorded throughout the experiment and for 5 min after the experiment (see Table 1). We acknowledge that our study did not include a separate sham-delay control to rule out the potential influence of the 5-min time gap alone. We recorded latency, behavioural frequency, emotional state, initial decision and final decision as described above.

2.9. Data analysis

The video-recordings were scored manually onto datasheets by JAH, and a randomly selected 15 % subset was independently scored by an observer naïve to treatment (inter-observer B = 0.98 – 0.99). All analyses were conducted in R 4.3.1 (2023–06–16 ucrt), using data from

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

Dependent variables	Method of calculation
Behavioural diversity	The number of unique behaviours per 10 s from postures listed in Table 2, that the horse displayed 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.
Latency (seconds)	Time taken from the point where the horse entered the maze to the point where the horse exited the maze. From this, the 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 and negative were calculated as per Table 2. A horse was said to be in a particular state using the criteria described in text and 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
Validated posture ethogram used for 10-s scan sampling and emotional-valence scoring (after Torcivia and McDonnell, 2021 = 1 and Waring, 2003 = 2).

Posture	Positive criterion	Negative criterion	Reference
Ears	Tips ≥ 30° forward	Flat back	1
Eyes	Lids soft, pupils normal	Wide, pupils dilated	1
Nostrils	Relaxed, no flare	Flared, rapid breathing	1
Head-neck	Level or slightly down	High/tense or snaking	1
Tail	Hanging or gently swishing	Clamped or lashing	1
Trunk	Aligned with neck	Twisted/away	2
Gait	Walk/trot, relaxed cadence	Stiff, high stepping	2
Muscle tone	Soft	Rigid	2

both replicate runs of each treatment. Normality was assessed with Shapiro-Wilk tests; no variable was normally distributed ($p < 0.050$). Outcome distributions were inspected visually (histograms, Q-Q plots) and with DHARMA simulation. Latency and behavioural-frequency variables were strongly right-skewed and were best modelled with a Gamma family and log link, whereas emotional-state counts followed a Poisson family (dispersion = 0.97). Alternative families (log-normal for latency; negative-binomial for counts) yielded higher AIC values, and DHARMA residual diagnostics showed no over-dispersion (all $p > 0.20$) and no zero-inflation. Final models were fitted with the glmmTMB and lme4 packages, residuals checked with DHARMA, and Wald χ^2 tests computed with the car package. In addition to null-hypothesis testing, we report paired-samples Hedges’ g (small-sample corrected) with 95 % CIs for all treatment contrasts, and we ran a post hoc sensitivity analysis (G*Power 3.1) to confirm the study’s detectable effect size given the within-subject design ($\alpha = 0.05$) (Faul et al., 2007).

Separate GLMMs were fitted for latency (pre- and post-decision), behavioural frequency (pre- and post-decision), and emotional-state counts; significant fixed effects were assessed with Tukey-adjusted pairwise contrasts of estimated marginal means using emmeans. Horse identity (HorseID) was fitted as a random intercept, and each baseline-treatment pair was assigned a unique SessionID nested within HorseID (random structure 1 | HorseID/SessionID) to account for repeated measures. The SessionID term explained ≤3 % of total variance (median SD = 0.07), and likelihood-ratio tests showed that its removal did not alter fixed-effect estimates (all $p > 0.50$); it was therefore retained for completeness. In addition, preliminary models that added Calendar Day or Repeat Number as additional random factors produced singular fits and accounted for <2 % of variance, so the more parsimonious random-effects structure was adopted.

Fixed effects were Treatment, Centre (intensive vs. extensive) and Sex, plus their *a priori* interactions Treatment × Centre and Treatment × Sex. The first interaction tested whether the acoustic-stress and object-play manipulations differed between intensively managed riding-school horses and extensively managed pasture horses, given contrasting baseline arousal and habituation to handling. The second addressed documented sex-specific stress responses and cognitive styles in horses (Wulf et al., 2013). Although neither interaction reached statistical significance (see Results), both were retained to avoid inflated type-I error rates (Zuur et al., 2009).

An initial decision was deemed successful (Y) if the horse chose the maze arm containing its preferred food and unsuccessful (N) if it chose the arm containing rocks; the same criterion was applied to the final decision. Because initial and final choices within a run were structurally dependent (YY, NN, YN, NY) and the resulting 2 × 2 table is sparse (68 paired observations), decision-success combinations were analysed with two-tailed Fisher’s exact tests, which are robust for sparse data; odds ratios are reported where significant (Agresti, 2013).

3. Results

3.1. Behavioural diversity

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 was highest in the stress treatment, followed by the play treatment (Fig. 2a). 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 of horses at the Centre B (intensively managed) was greater than those on the Centre A (extensively managed) Fig. 2b). There was a significant effect of sex on behavioural diversity before the decision point (GLMM: $\chi^2 = 9.44$, $df = 1$, $p = 0.002$), with behavioural diversity being greater in the males than females (Fig. 2c). 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 Centre B was greater than those of the horses at the Centre A in the baseline, stress and play treatments (Fig. 2d). There was no significant effect between treatment and sex on behavioural diversity (GLMM: $\chi^2 = 0.53$, $df = 2$, $p = 0.765$).

Centre 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 of the horses at Centre B was greater than those at Centre A (Fig. 3a). The interaction between treatment and Centre was also 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 of the horses at the Centre B was greater than those of the horses at the Centre A in the stress and play treatments (Fig. 3b). Treatment GLMM: $\chi^2 = 1.81$, $df = 2$, $p = 0.405$, sex (GLMM: $\chi^2 = 1.10$, $df = 2$, $p = 0.295$) and the interaction between treatment and sex (GLMM: $\chi^2 = 4.16$, $df = 2$, $p = 0.125$) were not a significant predictor of behavioural diversity after the decision point. 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.

3.2. Latency

Treatment was a significant predictor of latency to the decision point (GLMM: $\chi^2 = 8.15$, $df = 2$, $p = 0.017$), being shorter in the stress treatment than the play treatment (post hoc tests) (Fig. 4a). (post-hoc tests; Fig. 4a). Importantly, this shorter latency was **not** accompanied by better performance; rather, it reflected an impulsive, escape-motivated response that often resulted in an incorrect arm choice (i.e. shorter latency but poorer decision-making—quicker movement coupled with a higher error rate; see decision-success analysis below). Centre also had a significant effect on the latency before the decision point (GLMM: $\chi^2 = 6.80$, $df = 1$, $p = 0.009$), with horses on the Centre A reaching the decision point slower than that of the horses at the Centre B (Fig. 4b). There was no significant effect of sex (GLMM: $\chi^2 = 2.98$, $df = 1$, $p = 0.084$), the interaction between treatment and Centre (GLMM: $\chi^2 = 5.33$, $df = 2$, $p = 0.070$) and interaction between treatment and sex (GLMM: $\chi^2 = 1.53$, $df = 2$, $p = 0.466$) on latency before the decision point.

Centre had a significant influence on the latency after the decision to end of the maze (GLMM: $\chi^2 = 12.50$, $df = 1$, $p < 0.001$); the latency of horses at the Centre B was shorter than those on the Centre A. 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 Centre B was shorter than those at the Centre A in the stress treatment. There was no significant effect of treatment (GLMM: $\chi^2 = 1.31$, $df = 2$, $p = 0.519$), sex (GLMM: $\chi^2 = 0.33$, $df = 1$, $p = 0.565$), the interaction between treatment and sex (GLMM: $\chi^2 = 2.33$, $df = 2$, $p = 0.312$) on the latency after the decision point.

Overall, there was a positive relationship between the latency before and after the decision point ($R = 0.38$, $p < 0.001$), indicating that the

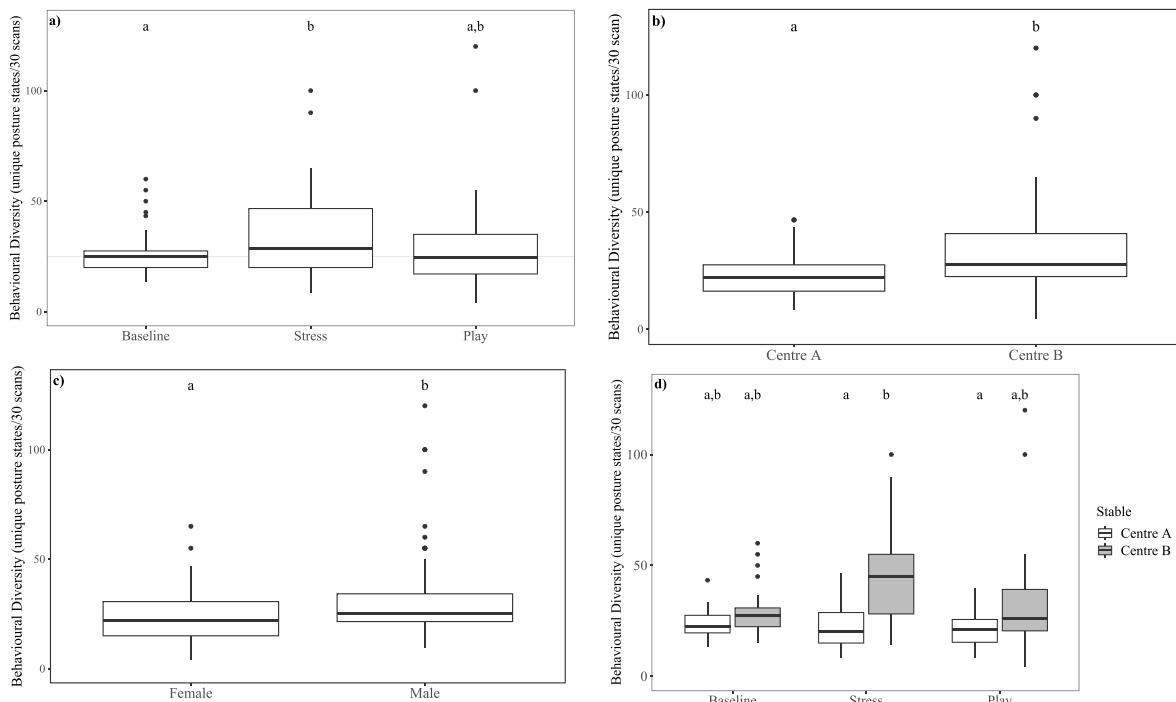


Fig. 2. Behavioural diversity before the decision point a) in three treatments, b) at two Centres (Centre A - extensively managed; Centre B - intensively managed), c) between female and male horses, d) interaction between treatment and stable. Boxplots indicate the median (central dark line), the 1st and 3rd interquartiles (boxes), minimum and maximum scores (whiskers) and outliers (dots). Bars with the same alphabets are not significantly different.

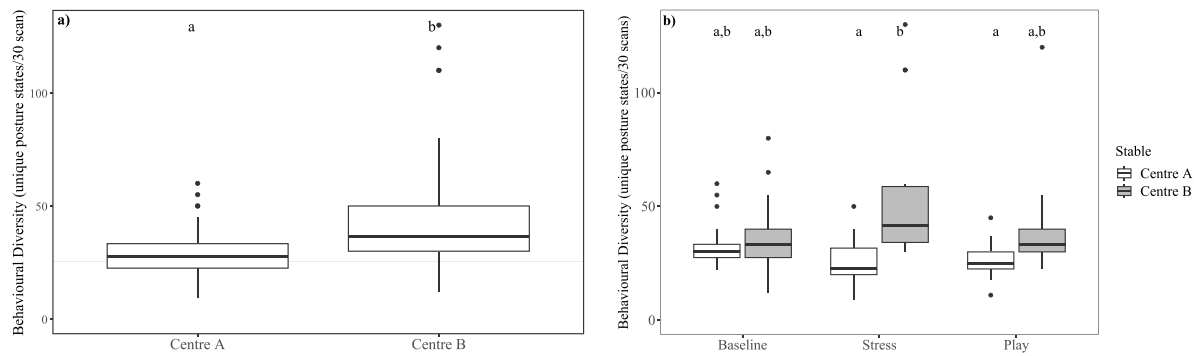


Fig. 3. Behavioural diversity after the decision point for a) at two stables, and b) at two Centres (Centre A - extensively managed; Centre B - intensively managed) in the three treatments. Boxplots indicate the median (central dark line), the 1st and 3rd interquartiles (boxes), minimum and maximum scores (whiskers) and outliers (dots). Bars with the same alphabets are not significantly different.

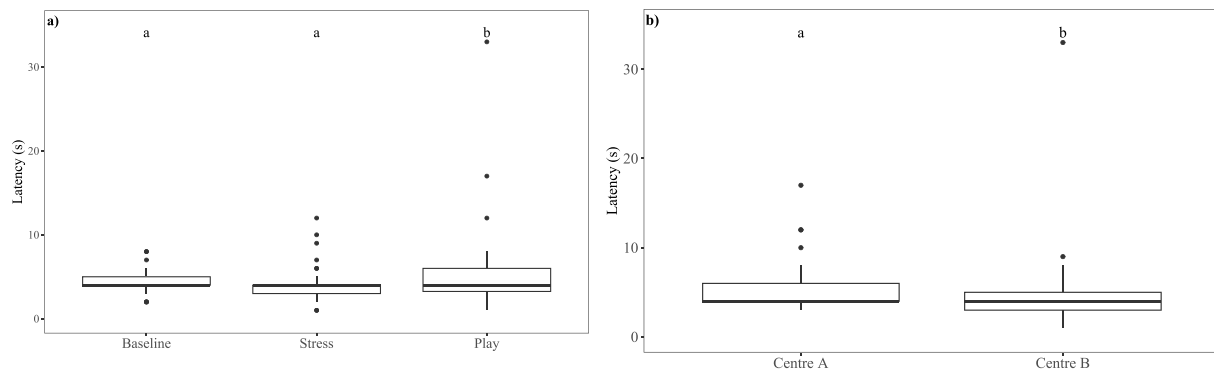


Fig. 4. The latency before the decision point a) in three treatments, b) at two Centres (Centre A - extensively managed; Centre B - intensively managed). Boxplots indicate the median (central dark line), the 1st and 3rd interquartiles (boxes), minimum and maximum scores (whiskers) and outliers (dots). Bars with the same alphabets are not significantly different.

latency of the horses was consistent when to the decision point and from the decision point to the exit of an arm.

3.3. Emotional states

Using the ear-eye-neck valence index, Treatment was a significant predictor of positive emotional state (GLMM: $\chi^2 = 9.03$, $df = 2$, $p = 0.011$), with positive emotional states occurring more frequently in the play treatment than the stress and treatment (Fig. 5a). There was no significant effect of stable (GLMM: $\chi^2 = 2.57$, $df = 1$, $p = 0.109$), sex (GLMM: $\chi^2 = 0.06$, $df = 1$, $p = 0.805$), the interaction between treatment and Centre (GLMM: $\chi^2 = 0.74$, $df = 2$, $p = 0.690$) and the interaction between treatment and sex (GLMM: $\chi^2 = 2.87$, $df = 2$, $p = 0.238$) on positive emotional state.

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 was greater than those in the baseline and play treatment (Fig. 5b). A significant effect was observed of Centre on negative emotional state (GLMM: $\chi^2 = 4.46$, $df = 1$, $p = 0.035$). The horses at the Centre B showed a greater level of negative emotional states than the than those on the Centre A (Fig. 5c). There was no significant effect of sex (GLMM: $\chi^2 = 0.01$, $df = 1$, $p = 0.942$), the interaction between treatment and Centre (GLMM: $\chi^2 = 3.72$, $df = 2$, $p = 0.155$) and the interaction between treatment and sex (GLMM: $\chi^2 = 1.37$, $df = 2$, $p = 0.503$) on negative emotional states.

3.4. Initial decision success and final decision success

An initial decision (Y = arm with preferred food; N = arm with

rocks) was recorded based on the horse's head position at the decision point; a final decision was scored in the same way. Across the four decision trajectories (Y-Y, Y-N, N-Y, N-N), there was no departure from chance in either the baseline or stress treatments: in the baseline condition horses chose Y-Y in 11 of 32 trials (34.4 %), Y-N in 10 of 32 (31.3 %), N-Y in 5 of 32 (15.6 %) and N-N in 6 of 32 (18.8 %; Fisher's exact test, $p = 0.260$), while under stress the corresponding frequencies were 2 of 31 (6.5 %), 8 of 31 (25.8 %), 10 of 31 (32.3 %) and 11 of 31 (35.5 %; $p = 0.283$). In contrast, the play treatment produced a highly skewed distribution: horses selected the preferred-food arm on both the initial and final decision in 21 of 22 trials (95.5 %) and switched from a correct to an incorrect arm in only one trial (4.5 %), with no N-Y or N-N outcomes, yielding a highly significant Fisher's exact test ($p < 0.001$); the odds of a consistent correct choice (Y-Y) were therefore 29.4 times greater than the odds of a correct-then-incorrect choice (Y-N).

3.5. Overall effect sizes and sensitivity

Beyond the model test statistics reported for each outcome, we examined effect sizes to gauge practical importance. For the primary endpoint (decision accuracy), the play-stress contrast for the initial choice yielded a moderate paired-samples Hedges' $g = 0.58$ (95 % CI: 0.15 – 1.02). A post hoc sensitivity analysis (G*Power 3.1, $\alpha = 0.05$, repeated-measures design) indicated 87 % power to detect an effect of this magnitude with 17 horses. The corresponding effect for final success was smaller and non-significant (Hedges' $g = 0.23$, 95 % CI: -0.04 – 0.49). For the GLMM outcomes, behavioural diversity, latency, and emotional state, the marginal R^2 (Nakagawa and Schielzeth, 2013) attributable to treatment ranged from 0.24 to 0.40, denoting

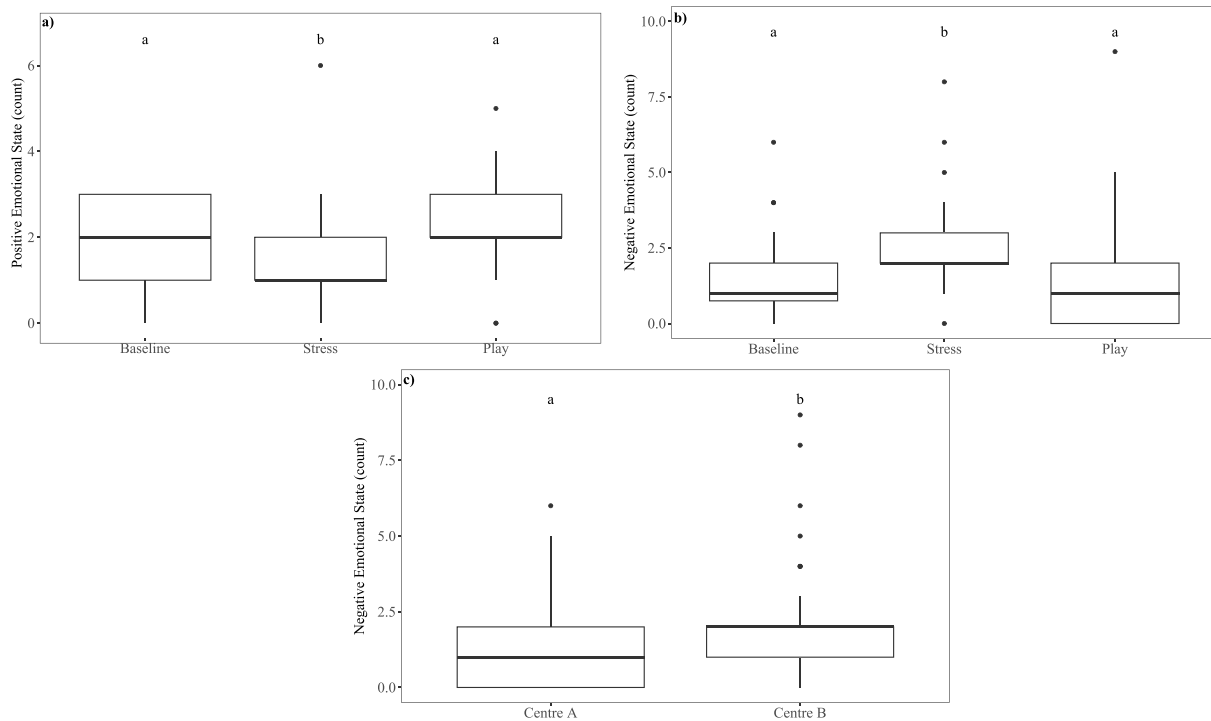


Fig. 5. The a) positive emotional state in the three treatments, b) negative emotional state in the three treatments, c) negative emotional state of the horses at two stables. Boxplots indicate the median (central dark line), the 1st and 3rd interquartiles (boxes), minimum and maximum scores (whiskers) and outliers (dots). Bars with the same alphabets are not significantly different.

medium-to-large effects, whereas interaction terms contributed less than 0.06. Taken together, these estimates showed that the observed treatment effects are statistically significant and behaviourally meaningful, despite the modest cohort size.

4. Discussion

We studied whether play buffers stress and improves decision-making in horses. We found support for our predictions that play improves emotional state and decision-making following a stressor. Contrary to our prediction, Centre A (extensively managed) horses were less affected by stressors compared to Centre B (intensively managed) horses.

When animals are exposed to new or stressful conditions, they may exhibit a wider range of behaviours to cope with these challenges. For example, Koolhaas et al. (1999) found that male rats displayed a wider variety of aggressive behaviours when given a mild shock. Similarly, Réale et al. (2007) found that some animals were bolder or more aggressive, displaying a wide range of behaviours in stressful situations. Based on these findings, the greater behavioural diversity in the stress treatment in our study indicates that the horses experienced a negative situation or stimulus, because behavioural diversity scores increased significantly from the baseline to the stress treatment. This outcome supports our initial assumption of greater behavioural diversity indicating a coping response to stress (Wechsler, 1995; Hall et al., 2018; Anderson et al., 2020). Moreover, the reduction of behavioural diversity in the play treatment, matching the baseline treatment, suggests that play buffers stress.

We used behavioural diversity scores to assess the range and variability of horse behaviours before the decision point in a Y maze. There is ongoing debate in the literature about whether behavioural diversity is a positive indicator of animal welfare. On one hand, low behavioural diversity indicates animal boredom or habitual behaviour (Rabin, 2003; Miller et al., 2020). On the other hand, higher behavioural diversity has been considered an indicator of good animal welfare, suggesting a

stimulating or enriched environment (Miller et al., 2020; Hall et al., 2021; Hall et al., 2018). For instance, bottlenose dolphins (*Tursiops truncatus*) with higher behavioural diversity had lower faecal cortisol levels, indicating reduced stress (Miller et al., 2021). However, our findings support the negative interpretation, showing that horses exhibited greater behavioural diversity under stress, likely as a coping response to negative stimuli, consistent with studies on rats and other animals (Koolhaas et al., 1999; Réale et al., 2007; Wechsler, 1995; Hall et al., 2018; Anderson et al., 2020). Conversely, the reduction of behavioural diversity during play treatment, returning to baseline levels, suggests that play alleviates stress, further complicating the interpretation of behavioural diversity as a welfare indicator.

The positive association of behavioural diversity 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 exposed to an elevated plus-maze continuously displayed a similar array of behaviours within the first 24 h in the maze (Andrade et al., 2003).

The horses moved to the decision point significantly faster in the stress treatment compared to the play treatment. Animals respond to stress in various ways. Some may slow down or speed up their movements and behaviours in a stressful situation (Baragli et al., 2014; Pellis and Pellis, 2013; Esch et al., 2019). For example, laboratory rats had a slower latency to enter a compartment when stressed (Metz et al., 2001). In contrast, stressed Arctic ground squirrels (*Urocitellus parryi*) sped up their movements in relation to a stressor (Boonstra, 2004). In our study, the horses responded more quickly in the stress treatment, possibly to distance themselves from a stressful situation. However, the responses in the baseline treatment were similar to the stress treatment. This may be because the baseline session was each horse's first exposure to the maze, and novelty is known to raise arousal in horses (Visser et al., 2002); unfamiliar settings and procedures can therefore induce a mild stress response that diminishes once the animals become accustomed to the test environment. In contrast, the horses did not move as quickly

through the maze in the play treatment.

A positive emotional state was significantly more common in the play treatment compared to the stress treatment, whereas a negative emotional state was significantly more prevalent in the stress treatment compared to 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., 2004). Lowland or spotted paca (*Cuniculus paca*) also displayed a positive emotional state during object play (Lima et al., 2024).

Our 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 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). Whether such a neuroendocrine stress response occurred in the stress treatment requires testing.

Play behaviours can enhance a stress response, as seen during social play in horses experiencing chronic stress (Hausberger et al., 2012). In contrast, play can reduce the stress response, as seen in laboratory rats (Pellis and Pellis, 2017). Neurobiologically, play can 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, buffering the effect of a stressor (DeVries et al., 2003; Wang and Aamodt, 2012). Whether object play similarly reduces stress is not known, but our study suggests that it might do so. Our data support the idea that brief object play can act as a coping strategy that rapidly down-regulates sympathetic arousal and facilitates parasympathetic recovery (Held and Špinková, 2011). Similar stress-buffering effects have been documented in adult horses engaging in self-soothing behaviours (Lesimple et al., 2010) and for foals offered manipulable enrichment items (Merkies and Franzin, 2021) and leisure horses managed in richer environments that exhibit lower hair-cortisol loads seasonally (Mazzola et al., 2021). The convergence of these studies suggests that allowing goal-free object interaction may be an efficient, low-cost method to restore cognitive flexibility after acute stress.

We tested the decision-making of the horses in the maze. In the play treatment, the horses made the correct initial and final decisions 29.4 times more often than the incorrect final decision. These values were significantly better than in the baseline and stress treatments. Stress can influence decision-making in animals physiologically. There is a link between the dysregulation of the HPA axis and cognitive bias, as seen in merino sheep (*Ovis aries*) (Verbeek et al., 2019). Stress also leads to animals making rash and incorrect or potentially harmful decisions (Starcke Brand, 2012; Barrett, 2015). This interpretation is supported by the shorter latency to the decision point shown by horses in the stress treatment: rapid movement through the maze appears to reflect impulsive, escape-motivated behaviour rather than superior cognition, and these rash choices frequently led to selection of the non-rewarded arm, lowering decision accuracy (Starcke Brand, 2012). By contrast, the modestly longer latencies recorded after play likely represent deliberative exploration; the extra sampling time allowed horses to gather more information and thereby improve accuracy (Chittka et al., 2009).

The association did not hold true for the baseline treatment, however, 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. 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 explain why the horses did not perform optimally during the baseline treatment, possibly because they were not experiencing any form of stress (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 responses in the play and baseline treatments, they were better able to make decisions.

The horses at Centre B (riding school) showed significantly greater behavioural diversity and shorter latencies both before and after the decision point than horses at Centre A (pasture managed), indicating higher arousal and stress reactivity in the more intensively managed environment. Consistent with this, Centre B horses also spent more time in a negative emotional state, possibly linked to their more frequent riding and confinement. Previous work also reports lower stress in horses kept under naturalistic conditions (Yarnell et al., 2015) and chronic stress in animals ridden daily by multiple riders (Hausberger et al., 2012). Confinement itself can increase boredom-related stress (Hausberger et al., 2012; Yarnell et al., 2015; Esch et al., 2019). The rapid latencies under the acoustic-stress treatment likely represent impulsive, escape-motivated movements whereas the slightly longer latencies recorded after object play probably reflect this additional sampling time improved accuracy, echoing classical speed-accuracy theory, as described above. Nevertheless, overall decision-making did not differ by housing system: there was no evidence that the less-constrained Centre A horses chose more accurately than the more-constrained Centre B horses.

Because breed identity was confounded with housing location and sample sizes were small, we could not include breed as an independent factor without over-fitting the models. Consequently, we cannot rule out that some differences in behavioural diversity or emotional reactivity reflected inherent breed or personality traits rather than stable management conditions. Future studies should use larger, within-breed samples or crossover designs to disentangle genetic and environmental influences.

Males in our study showed a broader behavioural repertoire during the 30 s that preceded the choice point than did mares. If behavioural diversity is interpreted as heightened arousal, this would suggest that males were the more-aroused sex. However, published evidence on sex differences in equine cognition is mixed. In a controlled battery of visual-discrimination and visuo-spatial tasks, Murphy et al. (2004) found that mares learned a colour cue faster than males, whereas geldings outperformed mares in one spatial-navigation test, indicating that the direction of the sex effect depends on the cognitive domain assessed. By contrast, large-scale survey data on ridden behaviour showed little support for sex-based stereotypes: mares were no more reactive than geldings across a range of ridden contexts (Aune et al., 2020). Consistent with that field observation, we detected no advantage for either sex in the accuracy of the final maze choice, suggesting that the transiently higher pre-decision arousal observed in geldings did not translate into enhanced or reduced decision-making once play had been introduced.

5. Conclusion

Our study demonstrated that play significantly improves the emotional state and decision-making abilities of horses following exposure to an acute stressor. This finding underscores the potential of play as a practical intervention to mitigate stress in horses, particularly those in high-stress environments such as riding schools. The observed stress responses to novel sounds confirm that horses perceive such stimuli as threats, triggering known stress behaviours. Importantly, the ability of play to return horses to a reduced stress state and enhance their decision-making suggests that incorporating play into their routine could be a valuable strategy for improving their welfare. The differences

in stress levels between horses from riding schools and those from pastures highlight the impact of stabling conditions on equine welfare. Horses in more constrained environments exhibited higher stress levels, indicating a need for better management practices in such settings. Overall, our findings advocate for the inclusion of play or other forms of enrichment in the management of horses to reduce stress and enhance cognitive function. By understanding and addressing the factors that contribute to stress in horses, we can improve their welfare and performance in various equestrian activities. Future studies should (i) add a brief sham-delay control to confirm that the effect is attributable to play rather than spontaneous recovery; (ii) validate the posture-based valence index with physiological measures (e.g. heart-rate variability, salivary cortisol) to strengthen objectivity; and (iii) recruit a larger, balanced, multi-centre sample to refine effect-size estimates and test breed-specific generality. Investigating longer-term mechanisms, such as parasympathetic activation or synaptic plasticity, would further illuminate how object play buffers acute stress.

CRedit authorship contribution statement

Jodi Anne Howard: Writing – original draft, Visualization, Formal analysis, Data curation, Conceptualization. **Neville Pillay:** Writing – review & editing, Writing – original draft, Visualization, Supervision, Funding acquisition, Conceptualization.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

Funding was provided by the University of the Witwatersrand. We are grateful to the owners of Centre A and Centre B for allowing us to use their horses and premises for this study. We thank Mr James Howard, Dr Duane Mol, Dr Therina Mol, and Ms Alexandra Viljoen for their technical assistance. Last and most certainly not the least, I (Jodi) need to give all the glory to God. I am because He is!

Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at [doi:10.1016/j.applanim.2025.106752](https://doi.org/10.1016/j.applanim.2025.106752).

References

- Agresti, A., 2013. *Categorical Data Analysis*, third ed. Wiley, Hoboken NJ.
- Anderson, C., von Keyserlingk, M.A.G., Lidfors, L.M., Weary, D.M., 2020. Anticipatory behaviour in animals: A critical review. *Anim. Welf.* 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. *Physiol. Behav.* 78, 125–133. [https://doi.org/10.1016/S0031-9384\(02\)00941-1](https://doi.org/10.1016/S0031-9384(02)00941-1).
- Aune, A., Fenner, K., Wilson, B., Cameron, E., McLean, A., McGreevy, P., 2020. Reported behavioural differences between geldings and mares challenge sex-driven stereotypes in ridden equine behaviour. *Animals* 10 (3), 414. <https://doi.org/10.3390/ani10030414>. PMID: 32131444; PMCID: PMC7142782.
- 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: Buss, D.M. (Ed.), *The Handbook of Evolutionary Psychology*. John Wiley & Sons, Ltd, pp. 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. *J. Equine Vet. Sci.* 40, 84–93. <https://doi.org/10.1016/j.jevs.2016.01.016>.
- Becker-Birck, M., Schmidt, A., Lasarzki, J., Aurich, J., Möstl, E., Aurich, C., 2013. Cortisol release and heart-rate variability in sport horses participating in equestrian competitions. *J. Vet. Behav.* 8, 87–94. <https://doi.org/10.1016/j.jveb.2012.05.002>.
- Bini de Lima, A.C., Sebastião da Fé, V.C., Palermo Hernandez, M.S., Oliveira dos Santos, V.M., 2023. Olfactory stimulation as environmental enrichment for domestic horses—a review. *Animals* 13, 3180. <https://doi.org/10.3390/ani13203180>.
- Boonstra, R., 2004. Coping with changing northern environments: the role of the stress axis in birds and mammals. *Integr. Comp. Biol.* 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. *Horm. Behav.* 43, 48–59. [https://doi.org/10.1016/S0018-506X\(02\)00022-3](https://doi.org/10.1016/S0018-506X(02)00022-3).
- Chittka, L., Skorupski, P., Raine, N.E., 2009. Speed-accuracy tradeoffs in animal decision making. *Trends Ecol. Evol.* 24, 400–407. <https://doi.org/10.1016/j.tree.2009.02.010>.
- Christensen, J.W., Rundgren, M., Olsson, K., 2006. Training methods for horses: habituation to a frightening stimulus. *Equine Vet. J.* 38, 439–443. <https://doi.org/10.2746/042516406778400574>.
- Christensen, J.W., Zharkikh, T., Chovaux, E., 2011. Object recognition and generalisation during habituation in horses. *Appl. Anim. Behav. Sci.* 129, 83–91. <https://doi.org/10.1016/j.applanim.2010.11.016>.
- Clark, F., Smith, L., 2013. Effect of a cognitive challenge device containing food and non-food rewards on chimpanzee well-being. *Am. J. Primatol.* 75. <https://doi.org/10.1002/ajp.22141>.
- Coleman, K., Novak, M.A., 2017. Environmental enrichment in the 21st century. *ILAR J.* 58, 295–307. <https://doi.org/10.1093/ilar/ilx008>.
- DeVries, A.C., Glasper, E.R., Detillion, C.E., 2003. Social modulation of stress responses. *Physiol. Behav.* 79, 399–407. [https://doi.org/10.1016/S0031-9384\(03\)00152-5](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>.
- Esch, L., Wöhr, C., Erhard, M., Krüger, K., 2019. Horses' (*Equus caballus*) laterality, stress hormones, and task related behaviour in innovative problem-solving. *Animals* 9, 265. <https://doi.org/10.3390/ani9050265>.
- Faul, F., Erdfelder, E., Lang, A.G., Buchner, A., 2007. G*Power 3: a flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behav. Res. Methods* 39, 175–191. <https://doi.org/10.3758/BF03193146>.
- 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>.
- Hall, C., Randle, H., Pearson, G., Preshaw, L., Waran, N., 2018. Assessing equine emotional state. *Appl. Anim. Behav. Sci.* 205, 183–193. <https://doi.org/10.1016/j.applanim.2018.03.006>.
- Hanggi, E.B., 2003. Discrimination learning based on relative size concepts in horses (*Equus caballus*). *Appl. Anim. Behav. Sci.* 83, 201–213. [https://doi.org/10.1016/S0168-1591\(03\)00136-9](https://doi.org/10.1016/S0168-1591(03)00136-9).
- 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. *Anim. Behav.* 81, 891–899. <https://doi.org/10.1016/j.anbehav.2011.01.007>.
- Kang, O.-D., Lee, W.-S., 2016. Changes in salivary cortisol concentration in horses during different types of exercise. *Asian Australas. J. Anim. Sci.* 29, 747–752. <https://doi.org/10.5713/ajas.16.0009>.
- Kirby, E.D., Muroy, S.E., Sun, W.G., Covarrubias, D., Leong, M.J., Barchas, L.A., Kaufer, D., 2013. Acute stress enhances adult rat hippocampal neurogenesis and activation of newborn neurons via secreted astrocytic FGF2. *eLife* 2, e00362. <https://doi.org/10.7554/eLife.00362>.
- 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. *Neurosci. Biobehav. Rev.* 35, 1291–1301. <https://doi.org/10.1016/j.neubiorev.2011.02.003>.
- Koolhaas, J.M., Korte, S.M., De Boer, S.F., Van Der Vegt, B.J., Van Reenen, C.G., Hopster, H., De Jong, I.C., Ruis, M.A.W., Blokhuis, H.J., 1999. Coping styles in animals: current status in behavior and stress-physiology. *Neurosci. Biobehav. Rev.* 23, 925–935. [https://doi.org/10.1016/S0149-7634\(99\)00026-3](https://doi.org/10.1016/S0149-7634(99)00026-3).
- Leiner, L., Fendt, M., 2011. Behavioural fear and heart rate responses of horses after exposure to novel objects: effects of habituation. *Appl. Anim. Behav. Sci.* 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>.
- Lundén, G., Oscarsson, R., Hedlund, L., Gjen, J., Jensen, P., 2022. Play ontogeny in young chickens is affected by domestication and early stress. *Scientific Reports* 12, 13576. <https://doi.org/10.1038/s41598-022-17617-x>.
- Lupien, S.J., Maheu, F., Tu, M., Fiocco, A., Schramek, T.E., 2007. The effects of stress and stress hormones on human cognition: implications for the field of brain and cognition. *Brain Cogn.* 65, 209–237. <https://doi.org/10.1016/j.bandc.2007.02.007>.
- Mazzola, S.M., Colombani, C., Pizzamiglio, G., Cannas, S., Palestini, C., Costa, E.D., Gazzonis, A.L., Bionda, A., Crepaldi, P., 2021. Do you think I am living well? A four-

- season hair cortisol analysis on leisure horses in different housing and management conditions. *Animals* 11, 2141. <https://doi.org/10.3390/ani11072141>.
- McEwen, B.S., Sapolsky, R.M., 1995. Stress and cognitive function. *Curr. Opin. Neurobiol.* 5, 205–216. [https://doi.org/10.1016/0959-4388\(95\)80028-X](https://doi.org/10.1016/0959-4388(95)80028-X).
- Meehan, C.L., Mench, J.A., 2007. The challenge of challenge: can problem-solving opportunities enhance animal welfare? *Appl. Anim. Behav. Sci.* 102, 246–261. <https://doi.org/10.1016/j.applanim.2006.05.029>.
- Mendl, M., Burman, O.H.P., Paul, E.S., 2010. An integrative and functional framework for the study of animal emotion and mood. *Proc. R. Soc. B Biol. Sci.* 277, 2895–2904. <https://doi.org/10.1098/rspb.2010.0303>.
- Merkies, K., Franzin, O., 2021. Enhanced understanding of horse-human interactions to optimize welfare. *Animals* 11, 1347. <https://doi.org/10.3390/ani11051347>.
- 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. *Physiol. Behav.* 72, 29–35. [https://doi.org/10.1016/S0031-9384\(00\)00371-1](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. *Behav. Neurosci.* 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>.
- Möstl, E., Palme, R., 2002. Hormones as indicators of stress. *Domest. Anim. Endocrinol.* 23, 67–74. [https://doi.org/10.1016/S0739-7240\(02\)00146-7](https://doi.org/10.1016/S0739-7240(02)00146-7).
- Murphy, J., Waldmann, T., Arkins, S., 2004. Sex differences in equine learning skills and visuo-spatial ability. *ISSN 0168-1591 Appl. Anim. Behav. Sci.* 87 (s1–2), 119–130. <https://doi.org/10.1016/j.applanim.2003.12.002>.
- Nakagawa, S., Schielzeth, H., 2013. A general and simple method for obtaining R^2 from generalized linear mixed-effects models. *Methods Ecol. Evol.* 4, 133–142. <https://doi.org/10.1111/j.2041-210x.2012.00261.x>.
- 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? *Learn. Behav.* 45, 355–366. <https://doi.org/10.3758/s13420-017-0264-3>.
- Potasz, C., De Varela, M.J., De Carvalho, L.C., Do Prado, L.F., Do Prado, G.F., 2013. Effect of play activities on hospitalized children's stress: a randomized clinical trial. *Scand. J. Occup. Ther.* 20, 71–79. <https://doi.org/10.3109/11038128.2012.729087>.
- 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. *J. Med. Bioeng.* 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. *Biol. Rev.* 82, 291–318. <https://doi.org/10.1111/j.1469-185X.2007.00010.x>.
- Sandi, C., 2013. Stress and cognition. *Wiley Interdiscip. Rev. Cogn. Sci.* 4, 245–261. <https://doi.org/10.1002/wcs.1222>.
- Sapolsky, R.M., 2002. Endocrinology of the stress-response. In: Becker, J.B., Breedlove, S.M., Crews, D. (Eds.), *Behavioural Endocrinology*, second ed. MIT Press, Cambridge, MA, pp. 409–450.
- Savin, H., Randle, H., 2013. Problem solving in the horse (*Equus caballus*) using a detour task. *J. Vet. Behav.* 8, e19–e20. <https://doi.org/10.1016/j.jveb.2012.12.044>.
- Selye, H., 1936. A syndrome produced by diverse nocuous agents. *Nature* 138, 32. <https://doi.org/10.1038/138032a0>.
- Shepherdson, D., 1994. The role of environmental enrichment in the captive breeding and reintroduction of endangered species. In: Olney, P.J.S., Mace, G.M., Feistner, A.T.C. (Eds.), *Creative Conservation: Interactive Management of Wild and Captive Animals*. Springer Netherlands, Dordrecht, pp. 167–177. https://doi.org/10.1007/978-94-011-0721-1_8.
- Shettleworth, S.J., 2001. Animal cognition and animal behaviour. *Anim. Behav.* 61, 277–286. <https://doi.org/10.1006/anbe.2000.1606>.
- Sigurdsson, T., Duvanci, S., 2016. Hippocampal-prefrontal interactions in cognition, behaviour and psychiatric disease. *Front. Syst. Neurosci.* 9, 190. <https://doi.org/10.3389/fnsys.2015.00190>.
- Starcke, K., Brand, M., 2012. Decision making under stress: a selective review. *Neurosci. Biobehav. Rev.* 36, 1228–1248. <https://doi.org/10.1016/j.neubiorev.2012.02.003>.
- Torcivia, C., McDonnell, S., 2021. Equine discomfort ethogram. *Animals* 11, 580. <https://doi.org/10.3390/ani11020580>.
- Valençon, M., Lévy, F., Prunier, A., Moussu, C., Calandreau, L., Lansade, L., 2013. Stress modulates instrumental learning performances in horses (*Equus caballus*) in interaction with temperament. *PLoS One* 8, e62324. <https://doi.org/10.1371/journal.pone.0062324>.
- van den Berg, C.L., Hol, T., Van Ree, J.M., Spruijt, B.M., Everts, H., Koolhaas, J.M., 1999. Play is indispensable for an adequate development of coping with social challenges in the rat. *Dev. Psychobiol.* 34, 129–138. [https://doi.org/10.1002/\(SICI\)1098-2302\(199903\)34:2<129::AID-DEV5>3.0.CO;2-U](https://doi.org/10.1002/(SICI)1098-2302(199903)34:2<129::AID-DEV5>3.0.CO;2-U).
- 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>.
- Visser, E.K., van Reenen, C.G., van der Werf, J.T., Schilder, M.B., Knaap, J.H., Barneveld, A., Blokhuis, H.J., 2002. Heart rate and heart rate variability during a novel object test and a handling test in young horses. *Physiol. Behav.* 76, 289–296. [https://doi.org/10.1016/S0031-9384\(02\)00698-4](https://doi.org/10.1016/S0031-9384(02)00698-4).
- Wang, S., Aamodt, S., 2012. Play, stress, and the learning brain. *Cerebrum* 2012, 1–9.
- Waring, G., 2003. *Horse Behaviour*, second ed. William Andrew Publishing, Noyes Publications, Norwich.
- Wechsler, B., 1995. Coping and coping strategies: a behavioural view. *Appl. Anim. Behav. Sci.* 43, 123–134. [https://doi.org/10.1016/0168-1591\(95\)00557-9](https://doi.org/10.1016/0168-1591(95)00557-9).
- Wulf, M., Aurich, J., May, A.-C., Aurich, C., 2013. Sex differences in the response of yearling horses to handling by unfamiliar humans. *J. Vet. Behav.* 8, 238–244. <https://doi.org/10.1016/j.jveb.2012.09.002>.
- Yarnell, K., Hall, C., Royle, C., Walker, S.L., 2015. Domesticated horses differ in their behavioural and physiological responses to isolated and group housing. *Physiol. Behav.* 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. *J. Comp. Neurol. Psychol.* 18, 459–482. <https://doi.org/10.1002/cne.920180503>.
- Zuur, A.F., Ieno, E.N., Walker, N., Saveliev, A.A., Smith, G.M., 2009. *Mixed Effects Models and Extensions in Ecology with R*. Springer, New York.