Title
The Comparative Effects of Three Enrichment Items on the Physiology and Behavior of Stalled Horses (Equus caballus)

Author
Brauns, Miranda

Publication Date
2024

Peer reviewed|Thesis/dissertation

The Comparative Effects of Three Enrichment Items on the Physiology and Behavior of Stalled Horses (*Equus caballus*)

By

MIRANDA BRAUNS
THESIS

Submitted in partial satisfaction of the requirements for the degree of

MASTER OF SCIENCE

in

Animal Biology

in the

OFFICE OF GRADUATE STUDIES

of the

UNIVERSITY OF CALIFORNIA

DAVIS

Approved:

Amy McLean, Chair

Kristina Horback

Maja Makagon

Committee in Charge

2024

Acknowledgements

To my Major Professor and committee chair, Dr. Amy McLean, you have been a joy to work with and I thank you so much for your wise mentorship, unwavering support, and incredible kindness throughout the entirety of my graduate program. I am truly grateful for all your time and effort; your constant guidance and assistance were invaluable to the completion of this thesis.

To my statistician, Dr. Ahmed Ali, thank you for your expertise and phenomenal work crunching the numbers. You have my sincere gratitude for your help with such an integral component of the thesis. Special thanks also go to Dr. Francisco Javier Navas González for his statistical consultation during study design.

To my committee, Dr. Kristina Horback and Dr. Maja Makagon, thank you for your guidance, support, and patience as I worked through everything. Your time, effort, and insight towards the thesis review are immensely appreciated.

To the UC Davis Equestrian Center facility managers, Holly Fox and Jaritza Granillo, the Equestrian Center residents, the ICA coaches, Jessie Weisinger and Jill Humphrey, and the NCEA team members, thank you for letting us use your facility and horses. Everyone was very considerate and excitedly interested during our research sessions (including the horses!), and I appreciate your time and collaboration very much.

To my incredible research assistants who volunteered and dedicated countless hours helping with study design, conducting research sessions, scoring behavioral data from video recordings, and transcribing datasheets, this project could not have been completed without you. Many thanks to Kavan Anderson, Abby Siegel, Lindsey Wright, Arden Uy, Julieta Herrera, Liat Bachar, Amber Ide, Michelle Green, Dellaney Rudolph-Gandy, and Chloe Bolanos.

To Dr. Jeannine Berger, thank you for sharing your expertise in horse research during the study design. Your insight was very helpful.

To Emily C. Dooley, thank you for writing an amazing UC Davis CAES media article. It was so exciting that you wanted to share our project.

To Dr. Cassandra Tucker, thank you for everything you've helped with during my graduate program. Your assistance and kindness have meant so much to me and are greatly appreciated.

Last but certainly not least, to my parents, Greg and Marcia, and my sisters, Hilary and Kristen, words cannot express how grateful I am for all your love and support. It is because of your guidance and encouragement that I know I can achieve whatever I set my heart and mind to. Thank you!

Abstract

Domestic horse management contrasts with their natural environment and behavioral needs. Small stalls and regulated feedings increase standing and reduce foraging which could impact the welfare of the horses. Enrichment seeks to encourage mental stimulation and species-typical activity budgets; however, more research is needed to understand how type of enrichment and provision time influences horse behavior. Furthermore, physiological parameters can provide insight of health and mental state, but enrichment effects on the physiology of horses are under-researched. This study analyzed physiology and behavior of nine stalled quarter horses throughout the day when provided a hay feeder (HF), activity ball (AB), and mirror (MIR) to better understand their effects on welfare. Horses were randomly divided into three groups and monitored over three trials, with each consisting of a control day and enrichment day. Each group received a different enrichment item per trial. Observations were conducted for 30 minutes four times per day (morning, noon, afternoon, and evening). Enrichment was removed between observations and during 5-day wash out periods between trials. Heart and respiration rate were recorded during observations approximately every 21 seconds through Nightwatch® Smart Halters™; behaviors were later scored via video recordings in BORIS using focal instantaneous scan sampling at 30-second intervals. Effects of type of enrichment, time-of-day, and possible interactions on each variable were tested using generalized linear mixed models; Tukey's HSD multiple comparison procedure were used for post-hoc comparisons (statistical significance set at $p \leq 0.05$). No significant differences in respiration rate were recorded, but enrichment significantly increased heart rate from control at all times (Morning: HF p=0.027, AB p=0.027, MIR p=0.022; Noon: HF p=0.022, AB p=0.028, MIR p=0.019; Afternoon: HF p=0.036, AB p=0.027, MIR p=0.024; Evening: HF p=0.038, AB p=0.031, MIR p=0.036). This indicates high arousal in response to enrichment, and heart rate increased the most when horses were provided HF versus other enrichment items. Provision time did not affect heart rate within any of the treatments except MIR in which horses showed lower heart rate in the evening (Morning p=0.026; Noon p=0.031; Afternoon

p=0.029). Between the three enrichment treatments, horses showed no significance difference in enrichment interaction and all items appeared equally engaging. Compared to control, all enrichments increased foraging (Morning: HF p=0.035, AB p=0.038, MIR p=0.041; Noon: HF p=0.029, AB p=0.036, MIR p=0.032; Afternoon: HF p=0.025, AB p=0.023, MIR p=0.039); decreased standing rest (Morning: HF p=0.036, AB p=0.035, MIR p=0.042; Noon: HF p=0.036, AB p=0.032, MIR p=0.041; Afternoon: HF p=0.023, AB p=0.026, MIR p=0.023); decreased standing alert (Noon: HF p=0.036, AB p=0.032, MIR p=0.037; Afternoon: HF p=0.029, AB p=0.028, MIR p=0.022); decreased social interactions (Noon: HF p=0.032, AB p=0.025, MIR p=0.039; Afternoon: HF p=0.021, AB p=0.029, MIR p=0.023); decreased frustration behaviors (Noon: HF p=0.035, AB p=0.041, MIR p=0.027; Afternoon: HF p=0.031, AB p=0.041, MIR p=0.043); and increased other behaviors (i.e., drinking, grooming; Afternoon: HF p=0.034, AB p=0.032, MIR p=0.029). However, while all items appeared effective enrichments overall, some variation in treatment effects were present. MIR resulted in no significant differences in locomotion compared to control; however, HF and AB increased locomotion (Morning: HF p=0.031, AB p=0.019; Noon: HF p=0.042, AB p=0.031; Afternoon: HF p=0.025, AB p=0.036). Horses showed increased foraging when provided HF at noon compared to AB and MIR (p=0.036 and 0.029, respectively) and had decreased standing alert in the afternoon compared to AB (p=0.036). HF and AB increased locomotion compared to MIR (Morning: HF p=0.039, AB p=0.043; Noon: HF p=0.026, AB p=0.023; Afternoon: HF p=0.032, AB p=0.037). Overall, the activity budget of horses when provided the HF suggests this item may more effectively fulfill overall behavioral needs; however, these results suggest that enrichment items could be selected to target specific behavioral needs and may be best implemented at certain times. Specifically, behavioral effects often appeared to be stronger during midday as seen within HF and AB treatments in which horses had increased locomotion at noon (HF: Morning p=0.039, Evening p=0.035; AB: Morning p=0.38, Evening p=0.026) and afternoon (HF: Morning p=0.034, Evening p=0.029; AB: Morning p=0.031, Evening p=0.036) and within MIR in which horses had increased enrichment interaction at noon (Morning p=0.029, Evening p=0.0369) and afternoon (Morning p=0.036, Evening p=0.038). Treatments also resulted in horses having a greater number of behavioral differences from control at noon and afternoon, compared to morning and evening. Therefore for better results, owners should prioritize giving enrichment when horses do not have routine meal provisions (likely midday). In summary, providing enrichment positively affects physiology and behavior of stalled horses which may improve horses' behavioral needs, health, and overall welfare. Further studies are recommended to assess different enrichment items, affective states, and long-term effects.

Table of Contents

List of Tables

List of Figures

1. Introduction

Free-roaming horses spend most of their time grazing, yet in captive management domesticated horses spend increased time standing stationary while grazing time is reduced by about 25% (Boyd et al., 1988; McGreevy, 2004; Ogilvie-Graham, 1994; Ransom & Cade, 2009). This compromises welfare by putting horses at heightened risk of developing ulcers, joint problems, boredom, and stereotypies (Gellman & Ruina, 2022; Harris, 1999; McGreevy et al., 1995). One reason for this predicament is that captivity conditions, such as small stalls and regulated feedings, do not properly align with the adaptations, biological needs, and behavioral motivations expressed by the animal as a result of the population's evolution in its natural environment (Fraser et al., 1997). Due to these limitations of captivity, enrichment is often given to improve welfare by providing opportunities that encourage mental stimulation and more species-typical activity budgets. In fact, several studies have shown that horses spend more time expressing foraging and locomotion behavior when their stall or bare paddock is enriched (Benhajali et al., 2009; Ellis et al., 2015; Goodwin et al., 2002; Hartman & Greening, 2019; Houpt et al., 2000; Huo, Wongkwanklom, et al., 2021; Rochais et al., 2018; Thorne et al., 2005). They also perform fewer stereotypic behaviors when provided with enrichment (Bulens et al., 2013; Grime et al., 2004; Henderson & Waran, 2001; Huo, Wongkwanklom, et al., 2021; McAfee et al., 2002; Mills & Davenport, 2002; Rochais et al., 2018). However, more research is needed to understand how type of enrichment item and provision time influences horse behavior. Furthermore, physiological parameters can provide insight of health and mental state, but enrichment effects on the physiology of horses are under-researched.

This study analyzed physiological and behavioral changes in stalled Quarter Horses when provided three enrichment items—hay feeder, activity ball, and mirror—each on different days. By comparing multiple enrichment items within the same study, we can better understand which behavior(s) each item influences and how each type of enrichment potentially affects horse activity budgets in stalled housing. Additionally, by conducting observations at four time points throughout the day, we seek to better understand how time-of-day influences enrichment use. Lastly, this study includes physiological analysis in addition to the behavioral component, something few studies have previously incorporated, outside music analyses (e.g. Kędzierski et al., 2017a; Kędzierski et al., 2017b; Neveux et al., 2016; Wiśniewska et al., 2019). The objective of the study is to better understand the effects of different enrichment items on horse behavior and physiology so we can more effectively utilize enrichment to improve horse health and welfare.

2. Literature Review

2.1 Horse Natural History, Domestication, & Management Practices

Horses (*Equus caballus*) are odd-toed ungulates that naturally live in plain lands across the world. Wild horses used to roam across most of the Western Hemisphere, but became locally extinct 10,000 years ago until the Spanish brought domesticated horses back to the West in the 1400's (McGreevy, 2004). According to the American Wild Horse Conservation, a nonprofit organization that manages and protects free-roaming horses and their habitat, feral horses are now found primarily on government-designated Herd Management Areas (HMAs) in ten western states: Arizona, California, Colorado, Idaho, Montana, Nevada, New Mexico, Oregon, Utah and Wyoming (2017). Feral, free-roaming horses live in large herds that typically travel around twenty miles per day and spend about 50–60% of their time grazing on grasses and other plants, spending only 20–35% of their time standing stationary (Boyd et al., 1988; McGreevy, 2004; Ransom & Cade, 2009). Foraging throughout the day is a particularly high motivated behavior in horses because they are adapted to a high fiber but low energy diet as a result of the population's evolution in its natural environment (Baumgartner et al., 2020). Due to the horse's strength, speed, and favorable traits for easy management, the species was an appealing candidate for domestication.

Horses were first domesticated in the Eurasian Steppes of Ukraine around 2000 B.C., with management changing based on the job required of the horse such as production, war, transportation, packing, performance, or companionship (McGreevy, 2004). Factors such as genetics, conformation, health and body condition, temperament, training, and behavior will aid the horse's ability to perform certain jobs well and thus determine for which purpose the horse will likely be used (McGreevy, 2004). Horses under human management are often viewed to have a certain value—economic, instrumental, and/or emotional—associated with their ability to perform their intended purpose (Endenburg, 1999; Posta et al., 2009; Tavernier, 1988). Depending on the horse's value and purpose, as well as the owner's attitude, knowledge, and available resources, husbandry management may differ between horse owners (Hemsworth et al., 2021). Many house their horses in large pastures or grassy paddocks (e.g., 150–450 m^2; Jørgensen & Bøe, 2007a) which provides space and opportunity to graze on grasses, move freely around the area, and socialize in small herds. For example, a Minnesota survey in which 659 respondents had an average of 14 horses reported that 87% of horses were kept on pasture during the summer (Martinson et al., 2006), a Canada survey with 76 respondents reported that 79% of the 211 horses were kept in outdoor group housing full-time (Ross et al., 2023), and an Australian survey with 57 respondents reported that 80% of the 98 horses were housed in paddocks full-time (Hemsworth et al., 2021). However, for management convenience or due to land ownership limitations, other owners instead keep their horses in small paddocks that may be barren of grass (e.g., 150–450 m^2; Jørgensen & Bøe, 2007a; 10 m x 30 m; Jørgensen & Bøe, 2007b) or individual stalls (typically 3.66 m^2 [12 ft^2] in the United States; Lewis, 2021). For example, a survey in Great Britain involving 873 horse owners reported the majority of horses were housed in a loose box/stable building or American barn (67.2% and 21.2%, respectively; Hotchkiss et al., 2007), a Canada survey with 76 respondents reported that 15% of the 211 horses were kept individually indoors (supplemented with time in outdoor group housing; Ross et al., 2023), and stalled housing has been noted to be particularly common for performance horses (Henderson, 2007; Werhahn et al., 2012). Stalls provide various benefits: horses are more easily accessible to the owners, there is more protection from weather, the amount and type of feed can be more easily regulated for individual horses, sick horses can be separated to prevent disease spread, and aggressive behaviors between horses can be reduced or completely prevented thereby reducing risk of injury and stress (McGreevy, 2004). In addition, owners may provide stalled horses with select time throughout the day in turnout pastures/paddocks, training, or going out for a ride. In the survey conducted by Hotchkiss et al. (2007), over 80% of owners reported providing full or partial day pasture access to their horses. Owners also reported exercising their horse for a median of six hours per week (Hotchkiss et al., 2007). However, in the survey conducted by Hemsworth et al. (2021), the most common duration of human-horse interaction time was less than 30 minutes per day and approximately 20% of the 57 participants did not engage in daily interaction with their horse. For horses in stalls, this limited interaction and restricted supplemental time outside their stalls could impact the horse's welfare.

2.2 Welfare Considerations

Welfare can be measured by an individual's physiological, psychological, and behavioral state given the conditions in which it lives (Dawkins, 1998; Fraser et al., 1997; Fraser, 2008). While people have shown concern about the appropriateness of captive environments for some time (Brambell, 1965; Shepherdson, 1998), animal welfare has gained great traction in recent decades (Perdue et al., 2020). In turn, standards for welfare continue to advance (Alligood & Leighty, 2015; Browning, 2020; Taylor et al., 2022). Enhancing welfare involves providing for the animal's physical and mental needs within the context of their behavioral biology and natural history (Broom, 1991; Dawkins, 1998; Fraser et al., 1997). To achieve this goal, animal husbandry management practices often refer to the "Five Freedoms" which states that animals deserve the right to live free from hunger and thirst; discomfort; pain, injury, or disease; fear and distress; and have the freedom to express normal behavior (Brambell, 1965; Webster, 2016). The ability to perform species-typical behavior is particularly imperative for animal development and learning appropriate skills needed for foraging, mating, and other fitness-related behaviors, as well as maintaining body condition, health, and positive mentality (McPhee & Carlstead, 2010). Therefore, "good welfare" aims to provide an animal's needs for healthy biological functioning as well as an environment in which the animal can perform species-specific behaviors (McPhee & Carlstead, 2010). Webster (2016) highlights a common concern, however, in that it is difficult to constitute what "normal behavior" entails, and that by assuming total control of the environment we deny animals the opportunity to promote their quality of life through behavioral choice. Wiepkema stressed this point back in 1987, stating that welfare problems occurred when animals are not able to control aspects of their environment. In this regard, the fifth freedom would be better expressed as "freedom of choice" (Webster, 2016). This phrasing incorporates more of a subjective understanding of welfare, an idea that is becoming increasingly more popular (Browning, 2020). Using a subjective approach, welfare is examined by the balance between positive and negative experiences of the animal, thus, "good welfare" should minimize negative experiences and promote positive experiences (Browning, 2020; McPhee & Carlstead, 2010). Overall, welfare is a complex objective that cannot be achieved or assessed by a single criterion and instead requires several comprehensive components pertaining to the physiological, psychological, and behavioral state of the animal (Broom, 1991; Dawkins, 1998; Fraser, 2008).

One reason for why the welfare of stalled horses may be at risk is due to the dissimilarities between stall management and the natural conditions in which horses evolved to live. This creates problems because some adaptations no longer serve a purpose in captivity (Fraser et al., 1997). Captive settings provide resources as well as protection from predators, disease, and bad weather, making behaviors such as aggression, antipredator and fear responses less important for the animal's survival (Fraser et al., 1997). As a result, domesticated animals often show modified behavioral needs compared to their wild counterparts and have higher or lower thresholds for behavioral response to stimuli (Price, 1984). This is due not only to the safety and resources of captivity, but also from breeders choosing more docile and less timid individuals for artificial selection and from animals associating the repeated exposure to humans with positive stimuli (Price, 1984). Though the need to express these kinds of behaviors may be reduced, they are still part of the natural repertoire that the animal may choose to act upon even in captivity.

Horses particularly are timid animals that are frequently observed to perform startle responses (i.e., defensive reflexes induced by unexpected stimuli which may involve body contractions, escape behavior, autonomic nervous system activation, and endocrine changes; Villas-Boas et al., 2016; Villas-Boas et al., 2022), even though they have been domesticated for thousands of years. In a study by Villas-Boas et al. (2016), horses performed startle responses to an abruptly opened umbrella with an average latency of 0.71 seconds, average duration of 3.97 seconds, and an escape distance of 3.43–9.96 m which involved a small jump followed by quick movement away from the umbrella. Subsequent research showed similar results and no significant differences in startle latency or duration between polo, dressage, and endurance horses, suggesting that startle behavior may be similar across horse breeds (Villas-Boas et al., 2022). Aside from the standard escape behavior involving jumping and displacement away from the stimulus, horses may also exhibit kicking and rearing during startle response (Villas-Boas et al., 2022). In fact, a survey of 265 people concluded that the leading cause of rider injury occurred from the horse being startled (27% out of 530 reported cases; Camargo et al., 2018). In addition to behavioral changes, horses showed increased heart rate and LF/HF ratio when exposed to the umbrella opening test (Villas-Boas et al., 2016; Villas-Boas et al., 2022). Dressage and endurance horses also showed an increase in cortisol levels 30 minutes after the startle stimulus (Villas-Boas et al., 2022). Therefore, while these reactive behaviors are natural, they appear to be associated with negative experiences and performing them impacts welfare by increasing risk of injury and stress for both the horse and the owner.

Stalled settings and husbandry practices also restrict the full, natural expression of adaptations which forces the horse to adjust its behavior. Studies show that welfare problems arise when an animal is not able to carry out behaviors they are motivated to perform due to environmental constraints or are unable to gain positive feelings from performing the behavior (Fraser et al., 1997; Goodwin, 2002; McPhee & Carlstead, 2010). Specifically, the small area of a stall considerably reduces locomotive behavior, and regulated feedings (typically 2 meals provided in the morning and evening of hay and/or quickly consumed high concentrate pellets) reduce overall foraging behavior and do not provide the opportunity for horses to graze throughout the entire day. Compared to horses in natural conditions, horses kept confined to stalls spend a larger portion of their time standing (40-58%) and only a small amount of their time eating high density food (36-47%; McGreevy, 2004; Ogilvie-Graham, 1994). This negatively impacts the health of the horse because increased time standing stationary increases strain and stiffness of joints and muscles which may lead to joint problems and poor body performance (Gellman & Ruina, 2022). At the same time, larger, less frequent meals with more concentrated food lead to faster ingestion, digestion, and empty gut fill, which increases risk of ulcers due to stomach acid splashing on walls of the empty stomach (Harris, 1999; McClure, 2016; Orsini, 2000). In addition, housing horses individually despite being a herd animal reduces opportunity to fulfill important social behavioral needs such as engaging in allogrooming, play, dominance displays, and avoidance behaviors necessary to forming social order which minimizes aggression among conspecifics (Goodwin, 2002; Søndergaard et al., 2011). This may be true even if the stalls allow visual and tactile contact between adjacently stalled horses due to the confined space limiting full expression of the social behavior. This is particularly important when considering that the inability to appropriately perform avoidance behaviors could increase risk of injury and stress to the horse, as supported by a study which found that 21.6% of the 231 injuries reported in a survey questionnaire were due to bites or kicks from conspecifics as a result of limited space (Knubben et al., 2008).

These behavioral modifications and reduced health may cause the horse to experience negative subjective feelings such as increased stress due to the reduced stimulation associated with the challenge of not being able to fulfill their behavioral needs or inability to cope with the environment (Fraser et al., 1997; McGreevy et al., 1995; McPhee & Carlstead, 2010). This is evidenced by a study in which horses kept in stalls and provided only 30 minutes of exercise per day for one month had significantly higher baseline cortisol levels on day 1 and day 29 of testing than those kept in paddocks (Park et al., 2013), and is supported by the fact that cortisol is commonly used to assess stress levels in many animal species (Karaer et al., 2023). Moreover, the threshold for fulfilling behavioral needs may increase due to the needs being modified through the domestication process (Price, 1984), making it even harder for owners to meet these needs and maintain welfare. As a result, some horses may perform anticipatory, restless, or frustration behaviors like pawing, kicking, and biting at stall components, and these behaviors may develop into undesirable abnormal behaviors called stereotypies that further impacts the animal's health and mental well-being (Cooper & McGreevy, 2002; McBride & Hemmings, 2009; McGreevy et al., 1995). Common stereotypies in horses include, but are not limited to: cribbing/crib-biting in which horses repetitively bite down on a horizontal surface and pull back with their teeth while contracting their neck muscles, horses may or may not suck in air; weaving in which horses repetitively sway their heads and neck side to side while shifting weight on each foreleg; stall/box walking in which horses repetitively walk around the stall in a repeated path and pattern; and repetitive self-biting, pawing, door-kicking, head-tossing, and wood chewing (Cooper & McGreevy, 2002; McBride & Hemmings, 2009; McGreevy et al., 1995).

2.3 Enrichment & Evaluations

With the aim to improve welfare, enrichment is a popular tool used among captive facilities. Enrichment is defined as a form of stimulation given to captive animals to promote positive experiences, provide opportunity and choice to perform motivated, species-appropriate behavior within the context of their behavioral biology and natural history, and/or mitigate unwanted behaviors that may otherwise harm the animal's well-being (McPhee & Carlstead, 2010; Newberry, 1995; Shepherdson, 1998). Through this opportunity, successful enrichment promotes improved biological functioning, cognitive thinking, and/or positive experiences (McPhee & Carlstead, 2010; Newberry, 1995; Shepherdson, 1998; Taylor et al., 2022). Enrichment can fall into non-mutually exclusive categories including but not limited to nutritional (e.g., slow feeders, scattered food), structural (e.g., climbing structures, substrates), sensory (e.g., scents, sounds), social (e.g. group composition), or cognitive (e.g., puzzles, toys, training), but all approaches seek to encourage species-typical behaviors and fulfill the animal's evolutionary need, or motivation, to carry out the behavior (Clark, 2017; Wells, 2009; Young, 2003). In other words, enrichment can help address and mitigate welfare concerns that arise when an animal's adaptations and needs do not correspond with the circumstances in captivity, as discussed in Fraser et al. (1997).

Although it is intuitively appealing to assume that enrichment will improve welfare of captive animals, our understanding of how and to what extent enrichment influences behavior is limited. To successfully utilize enrichment to improve welfare, it is crucial to consider several key variables. Firstly, is the enrichment biologically relevant to the species? Enrichment that lacks functional or goal-related significance to the animal will likely fail, as will enrichment based on incorrect causation hypotheses about a problem (Newberry, 1995). Secondly, which behavior(s) does the enrichment influence and does the enrichment produce the desired outcome? A puzzle feeder may be used with the intention to increase the duration of time an animal spends foraging; however, a study by Rochais et al. (2018) found that horses foraging from a hay bag hanging from a wall had increased frustration behaviors. Because this is not the desired outcome for this enrichment, use of that particular item may not be a suitable option for improving welfare. On the other hand, if the puzzle feeder does increase foraging without additional complications, it would then be an appropriate enrichment, as showcased by horses performing less frustration behaviors when foraging from a slow feeder on the ground compared to control and hay bag treatment (Rochais et al., 2018). Thirdly, when does the animal naturally express a behavior based on its circadian rhythms? If an animal typically rests in the shade midday to conserve energy from the heat, providing enrichment to promote locomotion at that time would not be suitable. To maximize welfare benefits, enrichment should be provided at opportune times of the day in conjunction with the animal's daily activity budget.

To ensure an enrichment's implementation is actually beneficial to the animal, it is important to evaluate enrichment use and effects. Subjective assessments in which conclusions are drawn through observation and personal judgment without use of formal research methodologies and statistical analyses can be biased or incorrect, thus systematic assessments should be used (Mellen & MacPhee, 2001). Newberry (1995) states that a reduction in abnormal behaviors could be an indicator of improved welfare, although this correlation is not always certain given that some abnormal behaviors may actually be adaptive in captive settings. Reduced negative emotional states could also be an indicator of improved welfare, but they cannot be easily measured directly (Newberry, 1995). On the other hand, improved health can be directly measured via physiological parameters and could be a good indicator (Newberry, 1995). To emphasis this concept, Newberry (1995) defines enrichment not as the item or environmental modification itself, but rather as an improvement in the animal's biological functioning as a result of the modification. However, in a literature review of enrichment evaluation, only 10 out of 94 studies (11%) included physiological measures (Alligood & Leighty, 2015).

Mellen and MacPhee (2001) introduce a five-point Likert scale that can be used for quick assessments which involves keepers scoring direct observational evidence of animals interacting with enrichment (1. flees, 2. ignores, 3. looks at with no contact, 4. brief contact, 5. repeated contact) and indirect evidence that the animal interacted with the enrichment outside of observations (1. no evidence, 2. moderate evidence, 3. significant evidence). Keepers can also assess whether the enrichment encouraged particular behaviors (0. undesirable behaviors, 1. no reaction, 2. reacted with behaviors unrelated to planned goal, 3. some goal behaviors, 4. moderate goal behaviors, 5. substantial amount of goal behaviors). However, this Likert scale method can only assess items added to an animal's exhibit and not the animal's entire environment; Mellen and MacPhee (2001) note that other measurement tools are needed for a more holistic approach to enrichment. More commonly, traditional animal behavior methodologies are used to evaluate enrichment, and the collected data is analyzed to look for changes in behavioral activity budgets, frequency of enrichment use, and response to an item over time (Alligood & Leighty, 2015; Mellen & MacPhee, 2001). Though it is important to note that inactive use of enrichment items does not mean welfare value is lost; enrichment can still reduce stress even when there is limited or no physical interaction for reasons such as the item offering opportunities for choice and control, some behavioral motivations being fulfilled quickly, and the stimuli being preferred as they signaled good environments to the animal's ancestors (Decker et al., 2023).

The idea of inactive, but still beneficial, use of enrichment adds value to the argument of using physiological measures as a welfare indicator because they could give keen insight into the subjective feelings, or emotional affective state, of the animal and indicate how the enrichment item is being perceived. Affective states are emotional 'moods', lasting minutes to hours, and can be differentiated based on arousal (high or low) and valence (positive or negative) (Russell, 1980). Physiological parameters such as heart rate can be used to identify arousal levels because autonomic-emotional responses are caused by the sympathetic nervous system (Endler et al., 1989). As an example, an increased heart rate could indicate high arousal (e.g., excitement or anxiousness) while a lower heart rate could indicate low arousal (e.g., bored or relaxed) (Horback & Parsons, 2019). However, physiological measures alone only indicate arousal level and cannot be used to determine valence (Horback & Parsons, 2019). Pairing physiological data with the behavioral context, on the other hand, could provide helpful clues as to how the enrichment is being perceived by the animal and estimate what the valence may be. For example, if an animal has increased heart rate in the presence of enrichment while displaying avoidance behaviors towards the item, the animal could be behaving this way out of anxiousness or fear (Leiner & Fendt, 2011), indicating a negative valence. This concept is supported by judgement bias testing, often conducted as go/no-go tests, in which animals positively or negatively respond to an ambiguous stimulus depending on their affective state (Horback & Parsons, 2019). While it is difficult to definitively and objectively conclude an animal's affective state (Horback & Parsons, 2019), these analyses provide insight as to what the affective state may be, as opposed to looking at physiological and behavioral measures independently, and can help researchers infer whether the enrichment is producing the intended positive effect to the animal's life.

Utilizing these evaluation methods, studies have shown increases in species-typical behaviors and decreases in undesirable behaviors when animals are exposed to enrichment, leading to improved overall health that indicates improved welfare (Newberry, 1995; Shepherdson, 1998; Taylor et al., 2022). However, it is difficult to successfully implement enrichment into husbandry management practices if we do not fully understand how the enrichment item influences physiology and behavior of the target species. This is evidenced by enrichment studies showing contrasting or inconclusive findings, as well as enrichment being underexamined for certain species, and this is particularly true in horses.

2.4 Horse Enrichment & Research Gaps

Popular enrichment items for horses include, but are not limited to, forage balls and hay nets, balls and barrels, scratching posts, mirrors, and music, and research has shown that these enrichments have many positive effects on horses. Most research on horse enrichment has studied foraging enrichment including items such as hay feeder balls, hay nets, and grated slow feeders, as well as spatially and/or temporally distributed food provisions (Benhajali et al., 2009; Ellis et al., 2015; Goodwin et al., 2002; Henderson & Waran, 2001; Jørgensen et al., 2011; Kenney, 2018; Rochais et al., 2018; Thorne et al., 2005; Winskill et al., 1996). Foraging enrichment has been reported to encourage appetitive and consummatory foraging and reduce time spent standing stationary (Benhajali et al., 2009; Ellis et al., 2015; Rochais et al., 2018; Thorne et al., 2005; Winskill et al., 1996). These results likely occur because the foraging devices increase the complexity and therefore time it takes for the horses to acquire the food. Appetitive and consummatory foraging also increased when horses were provided multiple foraging opportunities compared to single foraging opportunities (Ellis et al., 2015; Goodwin et al., 2002; Thorne et al., 2005). Additionally, horses have been shown to have more positive social interactions and less aggression when provided with foraging enrichment opportunities (Benhajali et al., 2009; Rochais et al., 2018). Due to the change in activity budgets and foraging needs being more appropriately met, the occurrence of frustration and stereotypic behaviors have also been shown to decrease when horses were presented a forage ball

(Henderson & Waran, 2001) and multiple forages (Goodwin et al., 2002; Thorne et al., 2005), although some studies found contrasting results that forage balls did not reduce stereotypic cribbing (Kenney, 2018) and horses showed item-directed frustration towards hay bags (Rochais et al., 2018).

While not as extensively researched, some studies have investigated the use of mirrors in horse stalls as it is thought the illusion of visual contact with another horse may be beneficial in minimizing social isolation (Grime et al., 2004; McAfee et al., 2002; Mills & Davenport, 2002). Horses have wide panoramic eyesight in front and on both sides, with binocular overlap oriented down the nose, and have a visual acuity of 20/33 able to best discern objects with contrast (McGreevy, 2004). These considerations provide evidence that horses are likely able to perceive mirror reflections due to the contrast of the images it is reflecting, particularly when there is movement via the horse or surroundings as this is something to which horses' eyes are extremely sensitive (McGreevy, 2004). Though clarity of the image's detail may depend on distance from the object, head angle, and environmental illumination (McGreevy, 2004). In fact, mirrors have been shown to significantly reduce stereotypic nodding (McAfee et al., 2002), weaving (Grime et al., 2004; McAfee et al., 2002; Mills & Davenport, 2002), and frustration behaviors like head shaking and door kicking (Grime et al., 2004). Reports of mirror effects on normal behaviors are contrasting, with Grime et al. (2004) seeing increased time ingesting hay and dozing during enrichment exposure compared to control, whereas McAfee et al. (2002) saw no effects on ingestion, dozing, or standing alert.

Another less commonly studied enrichment item includes ball-type toys which seek to encourage exploratory behavior (e.g., locomotion and play) by providing a toy that can be pushed around the stall (Bulens et al., 2015; Jørgensen et al., 2011; Kenney, 2018). Bulens et al. (2015) found that horses showed interest in Jolly balls (a 10" ball enrichment by Horseman's Pride made with durable material and a handle), but had limited item related behavior towards the items and they seemed to have no effect on normal or abnormal behaviors. Similarly, Kenney (2018) found that Jolly balls did not significantly reduce stereotypic cribbing.

While many studies analyze effects of a single enrichment item, it may be difficult to compare effects of different items across multiple studies due to the differences in study designs. Yet, limited research has been done comparing multiple enrichments in a singular study which is important as it offers insight for how to successfully strategize, implement, and evaluate enrichment husbandry management plans. Bulens et al. compared plastic sand-filled bottles and rope (2013) and compared rope to Jolly balls (2015), but horses in both studies showed little item-directed behaviors and enrichment did not have a significant effect on behaviors. Kenney (2018) compared a Jolly ball and feed dispenser but did not analyze normal behavior, instead focusing on stereotypic behavior and finding that neither enrichment had significant effects. In a more comprehensive comparison, Jørgensen et al. (2011) investigated the effectiveness of seven enrichment items in singly kept horses and four enrichment items in group kept horses during turnout. Items they investigated included a soft plastic cone, hard plastic ball, wooden scratching pole, peat soil, straw of barley, branches, and a concentrate feeder ball. They found that horses in both study groups directed more item-directed behavior towards edible objects, but found no relationship between item-directed behaviors and passive behaviors of standing and laying. However, access to straw reduced aggressive behavior in group kept horses (Jørgensen et al., 2011). In another study, Huo, Yaemklang, et al. (2021) compared a hay net, bedding straw, and yoga ball between horses that had stereotypies and those that did not. They found significant differences in the reactions to enrichment, with more rolling occurring when bedding straw was provided and more watching behaviors performed when the yoga ball was provided, indicating that different types of enrichments have differing effects. They also found differences in behavior between the stereotypic and non-stereotypic horses, indicating that such variables may influence enrichment use.

Due to the limited research and inconsistent findings involving non-edible enrichment items and comparison studies, the effects of enrichment on horses needs further investigation. We chose to compare horse enrichments commonly used by horse owners and that each represent a different sensory stimulus for a more comprehensive and unique comparison: a hay feeder (gustation), activity ball (tactile), and mirror (visual). Additionally, aside from studies which show evidence that music, specifically the genres of classical and country, reduces heart rate, heart rate viability, and cortisol levels indicating lowered stress levels (Kędzierski et al., 2017a; Kędzierski et al., 2017b; Neveux et al., 2016; Wiśniewska et al., 2019), very little research has been done on physiological effects of horse enrichment in addition to the behavioral component. Therefore, we chose to incorporate the physiological measurements of heart rate and respiration rate into our study to gain better insight into overall enrichment effects.

Lastly, behavioral schedules of horses and their effects on enrichment use is underexamined as well. Bulens et al. (2013) found that horses performed less item-directed behaviors towards sand-filled bottles and ropes in the afternoon and noon compared to the evening, but did not discuss if/how time of day had an effect on non-item related behaviors. Mills and Davenport (2002) found that when horses were provided a mirror, more alert and stereotypic behaviors were performed during the evening observation compared to the morning and midday observations. And when studying effects of a foodball, Winskill et al. (1996) found time-of-day effects on ingestion, movement, drinking, standing, lying, and nosing bedding; however, observations were conducted overnight so results are not able to give insight into enrichment effects during daytime. Therefore, we conducted research sessions at different times throughout the day and incorporated this variable into the analysis of our study to expand our knowledge about possible time-of-day effects on enrichment use.

3. Hypotheses & Predictions

We hypothesized that providing enrichment would affect heart rate, respiration rate, and behavior of stalled horses. Compared to control, heart and respiration rate were predicted to increase; foraging, locomotion, and enrichment interaction behaviors were predicted to increase; standing alert, standing rest, social, and frustration behaviors were predicted to decrease; and laying down, defecation, and other behaviors (e.g. grooming, drinking, urination) were predicted to have no change. We further hypothesized that type of enrichment would influence the physiology and behavior of horses to different degrees and there would be a time-of-day effect. We predicted the hay feeder would affect physiological and behavioral changes from control more than the activity ball and mirror, and the differences between enrichment treatments and control would be larger at noon and afternoon compared to morning and evening for all enrichment items.

4. Materials and Methods

4.1 Animals and Housing

A total of nine quarter horses, six geldings and three mares, with an average age of 11.33 ± 4.18 (mean ± SD) years (Table 1) and 5.5 body condition score (1-9 scale; Henneke et al., 1983) were included in this study. All horses were healthy and members of the National Collegiate Equestrian Association (NCEA) Division I Women's Equestrian Team. The horses observed in this study are managed in stall housing and experience similar durations of pasture/paddock access and training for jumping, reigning, hunter, and horsemanship performance horses, and are thus a good representation of a typical United States barn management schedule for performance horses. Results from this study are therefore relevant and applicable to the larger equine industry.

The horses were housed at the University of California Davis Equestrian Center in an open-air, center aisle style barn. All testing was performed at the Equestrian Center in each individual horse's stall. Stalls measured 3.6576 m. x 7.3152 m. (12 ft. x 24 ft.), divided into two equal sections: 1) a covered area with solid metal half walls (1.2192 m.; 4 ft. high) with bars on top (an additional 0.6096 m.; 2 ft. high), and 2) an uncovered area with metal bar half walls (1.397 m.; 4 ft. 7 in. high). The sliding stall door that divided the two sections remained open at all times. Rubber mat flooring was used for each stall and the covered area was additionally bedded with wood pellets and/or shavings. All fourteen stalls within the barn were occupied, including non-study individuals. Horses remained in their stall assigned via the Equestrian Center, so six study horses had two adjacent neighbors and three study horses housed in the end-cap stalls had one adjacent neighbor, with visual and tactile contact between conspecifics available through the stall bars.

Horses remained in their stalls during the full duration of each research session, and normal day-to-day husbandry management remained the same. Stalls were cleaned of waste daily and new bedding (minimum of 2 bags per stalls) was added to the covered stall areas weekly and as needed. Horses were provided three meals per day, the amount of which is calculated for individual horses at 1.5–2% of body weight: 1) grass hay in the morning (07300–0830 h), 2) a commercial grain diet (LMF Stable Mix or LMF Senior concentrate pellets) in the afternoon (1600–1900 h), and 3) alfalfa hay in the evening (1830–1930 h). Morning and evening hay provisions were provided on the ground. One horse of the nine was also given hay provision at noon (1200 h) every day in a hay net hanging about 0.46 m. (1 ft. 6 in.) off the ground. Concentrates were provided in a feed bucket at about 1.03 m. (3 ft. 4 in.) high off the ground. An automatic drinking water bowl was available in each stall, attached to the wall at a height of about 1.03 m. (3 ft. 4 in.) and water was accessible at all times.

Research was conducted in April 2022 with an average temperature of 15.5°C (59.9°F) and humidity of 42.6% (World Weather, 2022). All research days were sunny and partly cloudy.

4.2 Experimental Design

Horses were monitored over three trials, with each trial consisting of 2 days: a control day with no enrichment item and an enrichment day with one of the three tested items: hay feeder, activity ball, or mirror. Control days were always tested on day 1 and enrichment days were tested on day 2 for each trial. To minimize order effects of enrichment exposure, the nine horses were randomly divided into three groups of three, with each group receiving a different enrichment item per trial and rotating through the three enrichment items in subsequent trials (Table 1). By comparing data of an individual with and without enrichment, each horse served as its own control. Brand, color, and size information for the enrichment items are outlined in Table 2. Grass hay, separate from the daily meal provisions, was used to fill the hay feeders which was refilled to full capacity for each observation.

Horses were monitored for 30-minutes at four evenly spaced time points per day (0800 h, 1200 h, 1600 h, and 2000 h). The data collection schedule was chosen to reflect the variation across the different times of the day based on when the horses are typically in their stalls and the expected variation in the horse ethogram given the horse's daily management schedule (see section 4.1). Horses were equipped with a smart halter (Nightwatch® Smart Halter™, Protequus, Austin, Texas) 15-minutes before each time point to allow for desensitization of the halter before each observation started. During normal management practices, horses are taken out of the stalls for training or turnout immediately after halters are equipped, thus 15-minutes was determined as sufficient time for desensitization via previous personal observation. On enrichment days, the enrichment item was introduced at the top of the hour at each observation start, with the hay feeder and activity ball being placed in the center of the covered area of the stall and the mirror being hung on the metal bar of the front right wall of the stall with hooks and secured with tying rope. Both the Smart Halters™ and enrichment items were removed between each observation. Following the enrichment day, a 5 day (120 hour) wash out period was held between each trial. The study continued until all three enrichment items were tested for all three groups, spanning a total of three weeks.

Observations were recorded via video capture (Lorex 4K Ultra HD NVR and Security Cameras) from which behavior would be later scored. Cameras were setup along the ceiling of the center aisle of the barn, about 3.51 m. (11 ft. 6 in.) off the ground and pointed towards the stall. This setup allowed for full visibility of the entire stall, except for a small angle directly behind the solid half walls in the front of the stall and separating the covered and uncovered stall sections, putting the horse "out of view" only if its head was completely lowered in that area. Each horse was monitored by an individual camera. During the observation sessions of both control and enrichment days, the research team stood at the entrance of the barn, away from individual stalls as to not influence behavior from their constant presence. Every 5–10 minutes, one to two researchers walked through the barn to visually check all halters were still on and functioning (denoted by a small light on the left side of the halters behind the ear). The video capture was monitored during observations to ensure halter and horse safety, i.e., stop horses from excessively biting/pulling halters of neighboring horses, or if a horse showed adverse reactions to an enrichment item, then the team could remove the item for the rest of the session. Video capture was also used to check item status during enrichment days, i.e., if the hay feeder or activity balls rolled out of a stall at any point, the item was immediately retrieved and returned to the appropriate stall by placing it in the center of the covered stall area (as was done at the observation start). This occurred only with the activity ball and less than five times over all observations. Since the horses were already desensitized to people walking and talking through the barn as part of normal barn operations and behavior was recorded in intervals using scan sampling as described in section 4.3, these few moments of halter checks and enrichment replacements are assumed to have minimal or no observer effect on the data. This study was approved by the University of California's Institutional Animal Care and Use Committee (protocol #22748).

4.3 Dependent Variables

Recorded videos of the observations were scored using the Behavioral Observation Research Interactive Software (BORIS; version 7.13.6; Friard & Gamba, 2022). Behavioral data was recorded using focal animal, instantaneous scan sampling with 30-second intervals equating to 61 scans per observation per horse. Animal sampling was conducted by three observers, each being delegated one of the three trials for all nine horses. Details regarding training and inter-observer reliability are provided in section 4.4. Behaviors recorded and operational definitions are listed in the ethogram (Table 3).

Physiological data were auto-recorded around every 21 seconds from the Nightwatch® Smart Halters™. Outputs were given as beats per minute (heart rate) and breaths per minute (respiration rate) based on the average of the previous three successive readings using a moving average low pass filter. Physiological data equated to 86 data points per observation for each heart rate and respiration rate. Data was auto-uploaded to the associated mobile application every night when the halters were brought home and connected to Wi-Fi.

4.4 Statistical Analysis

To determine appropriate sample size, a power analysis based on the study parameters (number of dependent variables and observations) was conducted by Dr. Francisco Javier Navas González, Editor in Chief of *Zootechnics Archives* and academic researcher at University of Córdoba (Spain). Since our study involved a 1:2 ratio of parameters to observations per study individual, the Bayesian approach was used because it is more appropriate for evaluating smaller data sets while retaining power and precision (Hox et al., 2014; Lee & Song, 2004). With an alpha value of 0.01, calculations indicated a minimum of nine animals per treatment group was needed.

Statistical analyses were performed by Dr. Ahmed Ali, Assistant Professor at Clemson University, using R "stats" package software (version 3.3.1; R Core Team, 2013). The psych package was used to calculate descriptive statistics which were presented as mean ± standard error of the mean (SEM). Statistical significance was set as $p \leq 0.05$. The block study design of the trial setup in which there was a separate control day for each enrichment day and each enrichment day tested all items divided amongst the horses minimized possible day effects by allowing for an analysis of three control days versus three enrichment days. The effects of enrichment type, time-of-day, and all possible interactions on each variable (i.e., across treatments within times & within treatments across times) were evaluated using generalized linear mixed models (GLMM) with the lme4 package (Bates et al., 2014). Heart rate and respiration rate were first averaged for each 30-minute observation, then used in GLMM with the family set to "Poisson." Behaviors were calculated as percentages (number of occurrences of a specific behavior divided by the total number of observations at each time point multiplied by 100), then used in GLMM with the family set to “binomial.” Individual horses were entered in as a random effect for all models. Residual distribution and assumptions for GLMM of behavioral data were tested with the “DHARMa” package; normality of the model residuals for physiological data were evaluated with the Shapiro–Wilk test. Tukey’s honestly significant difference (HSD) multiple comparison procedure was conducted using the multcomp package (Hothorn et al., 2008) for post-hoc analyses of statistically significant effects of all models.

Inter-observer reliability was calculated during a 2-hour training period when the observers performed behavior video decoding on the same four, 15-minute videos that were not included in the current study. As the instructor, my own behavioral scoring of the same four videos was used as comparison for the agreement measure when calculating reliability. Cohen’s kappa agreement coefficient (κ) was used (Landis & Koch, 1977), and inter-observer agreement was considered good when Kappa exceeded 0.90. Observer reliability was reassessed at three evenly spaced time points throughout the data collection period to ensure reliability was maintained.

This study included 216 observations totaling 108 hours of video (12 hours per horse). Due to halter technical errors, 36.82% of physiological data was excluded from the analysis. This resulted from one halter acting faulty and missed the majority of data points from twenty observations for one horse. The remaining missing physiological data was limited per observation and were not from a single horse, rather the devices tended to skip a few readings on the 21 second scans. Behavioral data excluded from the analysis equated to 0.46% (one observation) due to a horse having a vet visit during a research session (hay feeder at noon) and 0.08% due to horses being “out of view.”

5. Results

5.1 Physiology

5.1.1 Heart Rate

The interaction between time-of-day and treatment for heart rate was significant (p=0.034). In the morning, afternoon, and evening, horses in the hay feeder treatment had higher heart rate compared to the activity ball, mirror, and control (Table 4; Figure 1). Horses in the activity ball and mirror treatments also had higher heart rate compared to control in the morning, afternoon, and evening; however, there was no difference between the activity ball and mirror treatments at these times. At noon, horses in the hay feeder and activity ball treatments had higher heart rate compared to the mirror and control, but there was no difference between the hay feeder and activity ball treatments. Horses in the mirror treatment also had higher heart rate compared to the control at noon.

Within the mirror treatment, horses had lower heart rate in the evening compared to the morning, noon, and afternoon (p=0.026, p=0.031, and p=0.029, respectively), but there were no differences between the morning, noon, and afternoon sessions (Table 4; Figure 1). There were no differences in heart rate within the control, hay feeder, and activity ball treatments across times of the day.

5.1.2 Respiration Rate

The interaction between time-of-day and treatment for respiration rate was not significant (Table 5; Figure 2). There were no differences in respiration rate across treatments or across times of the day.

5.2 Behavior

5.2.1 Foraging

The interaction between time-of-day and treatment for foraging was significant (p=0.032). During days when enrichment was not provided (control days), foraging was higher in the morning compared to noon and afternoon (p= 0.021 and 0.027, respectively) and was also higher in the evening compared to noon and afternoon (p= 0.023 and 0.026, respectively); however, there were no differences between morning and evening nor between noon and afternoon (Table 6; Figures 3–4). There were no differences in foraging within any of the enrichment treatments across times of the day.

Horses in the hay feeder, activity ball, and mirror treatments spent more time foraging in the morning and afternoon compared to the control, and there were no differences between the three enrichment treatments (Table 6; Figures 3–4). Horses in the hay feeder treatment spent more time foraging at noon compared to the activity ball, mirror, and control. Foraging in the activity ball and mirror treatments was also higher than control at noon, but there was no difference between the activity ball and mirror treatments. There were no differences in foraging between the hay feeder, activity ball, mirror, and control in the evening.

5.2.2 Locomotion

The interaction between time-of-day and treatment for locomotion was significant (p=0.036). When horses were provided the hay feeder and activity ball, locomotion was higher at noon compared to the morning (HF p=0.039; AB p=0.038) and evening (HF p=0.035; AB p=0.026). Locomotion was also higher within these treatments in the afternoon compared to morning (HF p=0.034; AB p=0.031) and evening (HF p=0.029; AB p=0.036). However, there were no differences in locomotion within these treatments between noon and afternoon nor between morning and evening (Table 6; Figures 3–4). There were no differences in locomotion within the mirror treatment or control across times of the day.

Horses in the hay feeder and activity ball treatments spent more time in locomotion in the morning, noon, and afternoon compared to the mirror and control, but there were no differences between the hay feeder and activity ball nor between the mirror and control (Table 6; Figures 3–4). There were no differences in locomotion between the hay feeder, activity ball, mirror, and control in the evening.

5.2.3 Standing Alert

The interaction between time-of-day and treatment for standing alert was significant (p=0.034). Horses in the hay feeder, activity ball, and mirror treatments spent less time standing alert at noon and afternoon compared to the control (Table 6; Figures 3–4). There were no differences between the three enrichment treatments at noon. Horses in the hay feeder treatment spent less time standing alert than the activity ball treatment in the afternoon, but there were no differences between the hay feeder and mirror nor between the activity ball and mirror. There were no differences in standing alert between the hay feeder, activity ball, mirror, and control in the morning and evening. There were no differences in standing alert within the control or any of the enrichment treatments across times of the day.

5.2.4 Standing Rest

The interaction between time-of-day and treatment for standing rest was significant (p=0.019). Horses in the hay feeder, activity ball, and mirror treatments spent less time standing rest in the morning, noon, and afternoon compared to the control, and there were no differences between the three enrichment treatments (Table 6; Figures 3–4). There were no differences in standing rest between the hay feeder, activity ball, mirror, and control in the evening. There were no differences in standing rest within the control or any of the enrichment treatments across times of the day.

5.2.5 Laying Down

The interaction between time-of-day and treatment for laying down was not significant (Table 6; Figures 3–4). There were no differences in laying down across treatments or across times of the day.

5.2.6 Defecation

The interaction between time-of-day and treatment for defecation was not significant (Table 6; Figures 3–4). There were no differences in defecation across treatments or across times of the day.

5.2.7 Social Behavior

The interaction between time-of-day and treatment for social behavior was significant (p=0.037). Horses in the hay feeder, activity ball, and mirror treatments spent less time engaging in social behavior at noon and afternoon compared to the control, and there were no differences between the three enrichment treatments (Table 6; Figures 3–4). There were no differences in social behavior between the hay feeder, activity ball, mirror, and control in the morning and evening. There were no differences in social behavior within the control or any of the enrichment treatments across times of the day.

5.2.8 Enrichment Interaction

The interaction between time-of-day and treatment for enrichment interaction was significant (p=0.036). When horses were provided the mirror, enrichment interaction was higher at noon compared to the morning and evening (p= 0.029 and 0.036, respectively) and higher in the afternoon compared to morning and evening (p= 0.036 and 0.038, respectively); however, there were no differences in enrichment interaction between noon and afternoon nor between morning and evening (Table 6; Figures 3–4). There were no differences in enrichment interaction within the hay feeder or activity ball treatments across times of the day. There were no differences in enrichment interaction between the three enrichment treatments at any of the time points.

5.2.9 Frustration Behavior

The interaction between time-of-day and treatment for frustration behavior was significant (p=0.013). Horses in the hay feeder, activity ball, and mirror treatments spent less time performing frustration behaviors at noon and afternoon compared to the control, and there were no differences between the three enrichment treatments (Table 6; Figures 3–4). There were no differences in frustration behaviors between the hay feeder, activity ball, mirror, and control in the morning and evening. There were no differences in frustration behaviors within the control or any of the enrichment treatments across times of the day.

5.2.10 Other Behavior

The interaction between time-of-day and treatment for other behavior was significant (p=0.023). Horses in the hay feeder, activity ball, and mirror treatments spent more time performing other behaviors in the afternoon compared to the control, and there were no differences between the three enrichment treatments (Table 6; Figures 3–4). There were no differences in other behaviors between the hay feeder, activity ball, mirror, and control in the morning, noon, and evening. There were no differences in other behaviors within the control or any of the enrichment treatments across times of the day.

6. Discussion

6.1 Control Group

Consistent and continuous foraging is a result of horses being adapted to a high fiber but low energy diet as a result of their population's evolution (Baumgartner et al., 2020); however, while horses in the control group foraged in the morning and evening at an average of 49.85% of the observations, foraging at noon and afternoon decreased to 30.56-31.58%. This is a result of the management schedule associated with regulated food provisions and presents welfare concerns because larger, less frequent meals which may include more concentrated food are known to lead to faster ingestion, digestion, and empty or less consistent gut fill in horses (Harris, 1999). Furthermore, while the amount of food ingested was not analyzed in this study to determine gut fill, horses had a large quantity of food from which to forage after receiving food provisions and only a low quantity of remains during midday times (personal observation) which suggests that the amount of food ingested during midday was likely not equally comparable to the morning and evening. This may increase risk of ulcer development because horses secrete gastric acid continuously (Campbell-Thomson & Merritt, 1987; McClure, 2016) and are therefore more susceptible to developing ulcers when they have less feed to neutralize the stomach acid (Harris, 1999; McClure, 2016; Orsini, 2000). This is evidenced by a study done by Murray and Schusser (1993) who recorded stomach pH levels in horses every six minutes over a 24-hour study and found that the percentage of pH readings less than two was significantly greater in unfed horses (hay withheld for 24 hours; 76% pH<2.0) than in fed horses (free access hay; 30% pH<2.0) and concluded that stomach acidity is highest when horses have less feed. Additionally, horses are left to fill the time gap with other behaviors. While not significantly different, horses within the control group had a higher percentage of standing alert, standing rest, social, and frustration behaviors at noon and afternoon compared to morning and evening. These tradeoffs could pose potential welfare risks. Specifically, standing stationary increases muscle and joint stiffness which can increase risk of joint pain and problems (Gellman & Ruina, 2022), and the performance of frustration behaviors could indicate stress or boredom, as evidenced by studies which show that frustration behaviors and stereotypies are often performed by horses and other animals when in a compromised or stressful environment (Garner, 2005; Leiner & Fendt, 2011).

We would like to note that the horses in our study were not known to perform stereotypic behavior and only performed frustration behaviors no more than an average of 11.96% of observations at a given time point. As described in section 4.1, study horses received husbandry practices in alignment with the Five Freedoms framework and were in good health. The welfare risks stated above (i.e., ulcers, injury, stress, stereotypies) are not an evaluation of the current state of the study horses, but rather intended to highlight possible risks that may occur in stalled horses in general, given the baseline results. Regardless of whether the starting point of welfare is considered "poor" or "good," the goal of enrichment is to enhance welfare by promoting positive experiences, providing opportunity to perform more species-typical behavior, and/or mitigate unwanted behaviors (McPhee & Carlstead, 2010; Newberry, 1995; Shepherdson, 1998). Thus, this discussion will focus on how the three tested enrichment items affect the horse's physiology and behavior with the goal to better understand if/how they can be effectively utilized to improve welfare.

6.2 Enrichment Treatments

6.2.1 Engagement with Enrichment

Horses engaged with all three enrichment items and there were no significant differences in enrichment interaction behavior between the enrichment treatments at all time points, so the items appeared to be equally engaging to the horses. In contrast, Jørgensen et al. (2011) found that horses kept individually in paddocks performed significantly more item directed behaviors towards edible items (13.2% of observation) compared to non-edible items (0.9% of observation). Moreover, Bulens et al. (2015) found that while Jolly balls appeared as equally interesting as ropes, both had limited item related behavior and had no effect on normal or abnormal behaviors of stalled horses. Bulens et al. concluded that nonedible items are not useful enrichment items for appropriately reared horses. On the other hand, the horses in our study engaged with the activity ball and mirror to similar proportions as the hay feeder, and all enrichment treatments had significant differences in heart rate and behavior from control. Thus, we argue that nonedible items can be useful as enrichment.

6.2.2 Physiological and Behavioral Effects

All three enrichment items significantly increased heart rate compared to the control at all time points. The increased heart rate could be the result of increased locomotion; however, locomotion only increased from control when horses were provided the hay feeder and activity ball in the morning, noon, and afternoon. This would not explain the increased heart rate recorded in the evening or when horses were provided the mirror. Additionally, when heart rate is increased due to increased activity, we would expect an associated increase in respiration rate because the body is working harder and needs more oxygen to handle the demand (European Respiratory Society, 2016), yet changes in respiration rate were not significant. This is supported by studies that show horses have locomotor-respiratory coupling at higher levels of activity (e.g., fast trot, canter, gallop); however, this coupling appears not to be observed with slower activity speeds (Art et al., 1990; Franklin et al., 2012). This could explain why our results of respiration rate were not significant, despite the increased locomotion, but supports our claim that the level of locomotion may not have been high enough to cause a change in heart rate. Therefore, activity is likely not the sole driver behind the increased heart rate. Rather, enrichment effects on heart rate could be due to the horse's emotional affective state. Specifically, increased heart rate is associated with high arousal (Horback & Parsons, 2019) and suggests that the hay feeder, activity ball, and mirror could be inducing affective states such as excitement (positive valence) or anxiousness (negative valence) within the horse. Similarly, Safryghin et al. (2019) found that horses with fewer behavioral responses to a pre-feeding experiment had higher heart rate and concluded that the heart rate was due to arousal rather than activity.

Horses had significantly higher heart rate when provided the hay feeder compared to the activity ball and mirror in the morning, afternoon, and evening and compared to the mirror at noon. This suggests that horses were in highest arousal state when provided with this enrichment item. This is likely due to the hay feeder providing a food reward, as grazing is a strong behavioral need in horses, and thus the horses were aroused to forage. This is supported by the fact that the hay feeder significantly increased foraging behavior compared to the activity ball and mirror at noon. The fact that the horses appear aroused to forage suggests that foraging behavioral needs are not being fully satisfied through the larger, less frequent meals of daily provisions typically provided in stall housing. This is the result of the regulated provisions contrasting with the horse's adaption of eating high fiber, low energy foods throughout the day with only short durations (three to four hours) of feed intake interruptions (Baumgartner et al., 2020). As such, Baumgartner et al. (2020) assessed the welfare of horses fed regulated meals 2–3 times a day and kept in stalls with edible versus non-edible bedding. They found that horses on non-edible shavings ate rations quicker, resulting in a longer period (9 hours overnight) of feed intake interruption (i.e., time between when roughage supply was available). Benhajali et al. (2009) found similar results in that horses given opportunity to forage spent significantly more time feeding compared to horses in a bare paddock (65.12%

vs. 29.75% of scans). Rochais et al. (2018) also found that horses significantly increased feeding time from hay bags (~40% of scans) and slow feeders (~45% of scans) compared to control hay on the ground (~28% of scans) and that the slow feeders also reduced undesirable behaviors. In addition, many horses were observed to engage in contrafreeloading during observations as many of them started foraging from the hay feeder immediately upon introduction, even if they were previously foraging from the daily meal provision on the ground, and other horses appeared to perform patch foraging by alternating their foraging between the hay feeder and daily provision (i.e., horses did not wait to finish their free access food before foraging from the hay feeder). In a study done by Winskill et al. (1996), horses directed foraging towards a foodball instead of *ad libitum* hay after eating concentrates, suggesting the foodball had a higher incentive value. Goodwin et al. (2002) found that horses that were provided multiple sources of forage showed significantly more foraging and significantly less alert behavior, locomotion, and eating of straw bedding compared to when forage was presented as a single source. Thorne et al. (2005), who also assessed single versus multiple forage diets, found that horses on multiple forage diet performed significantly more foraging behavior by an increase of about 20%, sampled all forages during the observation, and did not perform stereotypic weaving as the single forage horses did. While the diets in Goodwin et al. (2002) and Thorne et al. (2005) were free access, horses showing more species-typical foraging behavior when it was provided in multiple sources is supporting evidence for the benefits of contrafreeloading and patch foraging. The results of the aforementioned studies suggest that while horse nutritional needs may be met, the standard practice of regulated feedings compromise the behavioral and physiological aspects of welfare since it does not accommodate the horse's motivation for natural foraging behavior (Baumgartner et al., 2020), while the hay feeder enrichment seemed to help address such needs.

The increase in foraging when horses were provided the hay feeder could also be explained by the enrichment supplying more quantity of hay than the other treatments (feeder plus daily provisions) and simply having more to eat. However, the activity ball and mirror also increased foraging from control in the morning, noon, and afternoon, while showing no differences compared to the hay feeder in the morning and afternoon, despite those items not providing a food reward. If foraging was solely associated with food availability, we would expect a higher proportion of foraging in the morning and evening within the hay feeder treatment because they have "twice" the amount of food at those times (feeder + daily provision) compared to noon and afternoon in which they only have the food from the feeder available. Yet, horses spent almost equal percentages of scans foraging with the hay feeder at all time points, regardless of the quantity of food available at that time. This indicates that the increased foraging from control is not solely due to having more food to eat (hay feeder) and that other enrichment items can also stimulate the horse to engage in foraging behaviors.

The increase in foraging when horses are provided nonedible enrichments may be due to a novelty effect of the items, particularly considering that we removed enrichment between each observation session. In this case, we do not necessarily mean in the sense that the items themselves were novel to the horse but rather that the items are not usually or consistently present in the horse's stall under normal management. This novelty effect may have encouraged horses to investigate the area due to the novel objects or from the environment feeling changed. Conversely, McAfee et al. (2002) found that mirrors did not affect ingestion behavior and Bulens et al. (2015) stated that Jolly balls seemed to have no effect on normal behaviors.

Furthermore, the fact that there were significant differences in foraging within the control across times (i.e., increased foraging in the morning and evening), but not within enrichment treatments indicates that all enrichment items encouraged this highly motivated behavior at times when horses did not have routine meal provisions to the extent that foraging became more consistent throughout the day. This is beneficial for improving welfare as more consistent foraging should lower their risk of empty gut fill and ulcer development. While activity ball and mirror effects on the temporal distribution of foraging is under-researched, Rochais et al. (2018) found similar results to ours in that horses had a better temporal distribution of foraging when provided a slow feeder. They found that slow feeder horses always had hay available at mid-day and that 9.2% mean of horses still had hay left near the end of the day, while this almost never occurred with the hay bag and control group.

As with foraging behavior, locomotion also increased from control in the morning, noon, and afternoon when horses were provided the hay feeder and activity ball. Interestingly, rather than an increase in locomotion as we found with the hay feeder, Benhajali et al. (2009) found a decrease in locomotion when horses were given foraging opportunity compared to control (11.70% v. 23.56% of scans). The difference in our results could be due to their experimental area being a paddock with bigger space compared to the stalls in our study. While the horses that received enrichment in both studies had similar percentages of locomotion (11.70% in Benhajali et al. vs. 10.52% hay feeder and 11.85% activity ball in the current study), the control horses in the study by Benhajali et al. had more locomotion (23.56% of scans) than the control horses in our study (2.56% of the observation), and this influenced the direction of the behavioral change (decrease vs. increase). Winskill et al. (1996) also found that a foodball significantly decreased locomotion compared to baseline and observations were conducted in a stable; however, Winskill et al. notes this finding as an anomaly and deduced the reason being their ethogram incorporating both moving and eating while interacting with the foodball as one category. Thus, the increase in locomotion from our findings despite the small stall size provides evidence of these enrichments' positive effects. The increase in locomotion is likely due to both items needing to be moved for foraging, with the hay feeder needing to be pushed and manipulated to access the food and the activity ball needing to be moved out of the way of the daily provisions on the ground. In addition to the horse moving the activity ball for foraging, it created a larger obstacle in the stall that horses need to walk around when locomoting normally. It could be argued that this reduces the already limited space available within the stall, leaving the welfare benefits to be counterproductive; however, horses provided the activity ball show species-typical activity budgets.

Furthermore, the reduced frustration behaviors from control suggest that the activity ball likely did not add stress to the horses, so we conclude that the enrichment effects on welfare are still ultimately positive.

In conjunction with the increased foraging and locomotion, horses in all three enrichment treatments performed significantly decreased standing rest in the morning, noon, and afternoon as well as standing alert, social, and frustration behaviors at noon and afternoon compared to control. This is likely due to the significant increase in foraging and locomotion creating a tradeoff with the other behaviors. This is supported by Jørgensen et al. (2011) who found that for four of their seven item treatments, including a concentrate ball and toy ball, standing was significantly reduced when the turnout area had more roughage and horses spent more time eating. Similarly, Winskill et al. (1996) found that a foodball significantly decreased standing behavior compared to baseline and Benhajali et al. (2009) found that compared to horses in bare paddocks, horses given opportunity to forage spent significantly less time standing alert (5.23% v. 14.71% of scans) and standing rest (11.76 v. 27.52% of scans). On the other hand, McAfee et al. (2002) found that mirrors did not affect ingestion or time spent standing active. While this finding contrasts with ours, it may indeed suggest that the decreased standing observed when horses were provided a mirror in our study is primarily due to a behavioral tradeoff with foraging, rather than an independent effect the enrichment has on standing behavior. A tradeoff could also be occurring in response to filling more of their time interacting with the enrichment items. This finding contradicts Jørgensen et al. (2011) who found that item directed behaviors towards a concentrate ball and toy ball, among other tested items, did not reduce the occurrence of passive behaviors such as standing; however, their study observed horses in turnout areas and within group horses so the larger space and herd size may explain the differences in behavior observed between our studies.

Regardless of the tradeoff being caused by increased behaviors or interaction with the enrichment, our results of decreased standing, social, and frustration behaviors provide evidence of improved welfare from the potential concerns associated with the control. Firstly, since standing stationary increases muscle and joint stiffness (Gellman & Ruina, 2022), the increased locomotion and reduction in standing behavior should lower risk of joint pain and problems. Furthermore, the fact that enrichment affected both forms of standing behavior (i.e., alert and rest) suggests that enrichment was not only beneficial in encouraging the horse to be more active, but also in reducing the horses' desire and need to be vigilant possibly due to feeling more at ease in their environment associated with having better fulfillment of their behavioral needs. Secondly, while a decrease in social behavior initially seems like a negative effect for a social herd animal, we noticed that social behavior often appeared agonistic in control (personal observation) and suggest the reduction seen with enrichment may be a positive effect. Similarly, Benhajali et al. (2009) found that mares given access to forage performed fewer social interactions compared to control mares, and of the social behaviors observed, the experimental mares displayed significantly more affiliative behaviors (0.79 v. 0.24 frequencies per mare per hour) and less aggression (8.51 v. 15.42 frequencies per mare per hour). Jørgensen et al. (2011) also found that having access to straw in the turnout of group horses significantly reduced agonistic behaviors by about half, and though it was not significantly different among treatments, the straw treatment showed the highest percentage of affiliative behaviors. However, the type of social behavior was not recorded in the current study and therefore not statistically tested; more research needs to be done to analyze enrichment effects on affiliative versus agonistic social behaviors for a more comprehensive understanding. Lastly, since frustration behaviors have been known to be performed by horses and other animals when in a compromised or stressful environment (Garner, 2005; Leiner & Fendt, 2011), the reduced frustration behaviors when horses received enrichment treatments could indicate lowered stress and/or boredom. This is a positive effect on welfare not only for reducing stress and improving mentality, but because frustration behaviors performed as a result of unfulfilled behavioral needs can become repetitive and may contribute to the development of stereotypies (Cooper & McGreevy, 2002; Garner, 2005; McBride & Hemmings, 2009; McGreevy et al., 1995); therefore, the reduction in frustration behaviors suggests that enrichment may help mitigate stereotypic behaviors from developing. While stereotypic behaviors were not analyzed in the current study, Kenney (2018) found that neither a Jolly ball nor feed dispenser reduced cribbing behavior, other studies have found that certain enrichments do reduce stereotypy in horses. Mirrors significantly reduced stereotypic weaving compared to control (Grime et al., 2004; McAfee et al., 2002; Mills & Davenport, 2002); a sand filled bottle and rope had a tendency to reduce abnormal behaviors of licking walls and objects (Bulens et al., 2013); and a foraging device tended to reduce stereotypic behaviors (Henderson & Waran, 2001). Henderson and Waran (2001) similarly concluded that enrichment may help create an environment less likely to lead to the development of abnormal behaviors when used in conjunction with other measures. Further research utilizing a sample size of horses with stereotypies is needed to better understand enrichment effects on these behaviors.

6.2.3 Time-of-Day Effects

No significant differences were found between control and any of the enrichment treatments in the evening for any of the behaviors. This suggests that enrichment should likely be implemented during the day. This seems especially true for the mirror enrichment in which horses performed significantly lower enrichment interaction in the morning and evening compared to noon and afternoon. These results contradict Bulens et al. (2013) who found that horses displayed more item related behaviors towards sand-filled bottles and rope in the evening and Jørgensen et al. (2010) who found that item interest towards a concentrate ball and toy ball, among other tested items, was lowest in the afternoon. Our results are likely due to the horses having daily meal provisions in the morning and evening, causing them to be less interested in engaging with the mirror. Similarly, Bulens et al. (2013) found that when provided a sand filled bottle and rope, horses showed higher frequency of item related behavior when no hay was available. On the other hand, enrichment interaction within the hay feeder and activity ball treatments showed no differences across time. This is likely due to the fact that both items needed to be pushed for foraging in the morning and evening, as previously discussed, while also being engaging during midday when the environment was otherwise barren. In addition, the percentage of interaction with the mirror was lowest in the evening, and horses had a significantly lower heart rate with the mirror in the evening compared to the other times. The limited reflection that could be seen from the mirror in the darker light conditions may have caused the reduction in the mirror's physiological and behavioral effects. Similarly, Mills and Davenport (2002) found that horses provided a mirror performed more alert and stereotypic behaviors during the evening compared to morning and midday observations. Considering these two variables (darkness and meal provisions), we recommend providing enrichment—particularly mirrors—during daylight hours for more effective welfare benefits.

However, our study did not analyze the possible impacts of multiple exposure to enrichment, so the horses' behaviors may have been influenced by the evening session being the fourth time exposed to the enrichment on a single day. This consideration suggests that the lack of behavioral significance in the evening for all treatments may be due to the effects of enrichment dissipating over time with repeated exposure. Though, no significant differences in enrichment interaction were found within the hay feeder or activity ball treatments, indicating that horses engaged with those items in the evening at similar proportions to other times of the day. Moreover, while the behaviors of foraging and standing rest in the three enrichment treatments as well as locomotion in the hay feeder and activity ball treatments were significantly different from control in the morning, noon, and afternoon, many behaviors in the enrichment treatments (i.e., standing alert, social, frustration, and other behaviors) were only significantly different from control at noon and/or afternoon. If repeated exposure to enrichment did influence behavior, we would expect more behavioral changes during the first exposure in the morning. Thus, the results showing a greater number of behaviors being significant during midday suggests that there was a time-of-day effect in which providing enrichment, regardless of type, when horses do not have routine meal provisions (e.g., midday) appears to provide more positive behavioral effects on welfare. This is further supported by the results within the hay feeder and activity ball treatments in which there was significantly more locomotion at noon and afternoon compared to the morning and evening. The reason for this is unclear, but this finding is beneficial since locomotion in the control showed lowest percentages of locomotion at those times. As mentioned, standing stationary increases muscle and joint stiffness which can increase risk of joint pain and problems (Gellman & Ruina, 2022), so the increased locomotion at times when horses are not occupied fulfilling higher priority behavioral needs (e.g., daily meal provisions) should help keep joints loose and lower risk of joint injuries. Therefore, we recommend providing hay feeders or activity balls when horses do not have meal provisions to more effectively target locomotion behavioral changes for improved welfare. Winskill et al. (1996) also found that time-of-day influenced enrichment effects for all behaviors except eliminate when testing effects of a foodball, however they found peak durations of foodball interaction at sunrise and sunset. Though the contrast from our results is likely due to their study observing horses overnight as opposed to during the day as in the current study. Still, the results from Winskill et al. (1996) support our conclusion that enrichment seems to have more effective benefits when horses are not otherwise preoccupied fulfilling higher incentive behavioral needs (e.g., daily meal provisions, sleep, training/rides with owners).

Repeated exposure and long-term effects of enrichment are still important to consider, however, in which case providing different enrichments for shorter amounts of time but on a rotational basis may prove beneficial for welfare. This may be true as Jørgensen et al. (2011), who studied behavior during the first hour of turnout and the last hour of turnout with the enrichment item available throughout the entire turnout period, found that item directed behaviors in individually kept horses were significantly higher after the item was introduced indicating that horses lost interest over time. However, several studies have found evidence that enrichment can elicit effects over long-term periods, suggesting that novelty may not be entirely necessary for enrichment to maintain its welfare benefits. For example, Jørgensen et al. (2011) found that group kept horses did not lose interest in enrichment items across four days. While this contradicts the findings stated above that individually kept horses lost interest over time, they suggested there may be an aspect of social facilitation on enrichment use or could be due to the different study times between the two studies (individually kept horses were studied across one day per enrichment while group kept horses were studied only before noon and on days one and four with the same enrichment). These findings from Jørgensen et al. (2011) suggest that while enrichment use may be influenced by time-of-day based on the horse's daily activity budget, as similarly observed in our current study, horses do not appear to lose interest in enrichment across days. Thorne et al. (2005) also found that the effects of a multiple forage diet persist for a 7-day period. McAfee et al. (2002) found that effects of a mirror on reducing stereotypic weaving did not seem to dissipate over the 5-week treatment period. Lastly, Bulens et al. (2013) found that item related behavior towards sand filled bottles and rope did not significantly reduce after a week. However, Bulens et al. stated that the items aroused interest to a limited extent, with item related behavior occurring only an average of 3.85% of the observations, unlike the higher interaction seen in our study of an average of 14.56%. More research is recommended to further analyze dissipating effects as a result of repeated exposure and evaluate long-term benefits of enrichment, as well as how those effects compare across different enrichment items.

6.2.4 Overall Evaluation

As we previously defined, enhancing welfare involves caring for an animal's physical and mental needs within the context of their behavioral biology and natural history (Broom, 1991; Dawkins, 1998; Fraser et al., 1997) by providing an environment that promotes positive experiences and provides opportunity to perform species-typical behaviors (Browning, 2020; McPhee & Carlstead, 2010). As such, enrichment should only be evaluated as successful and effective at improving welfare if it helps achieve those objectives. Leiner and Fendt (2011) studied novel object tests in horses and found that physiological and behavioral fear responses are positively correlated, meaning that an increased heart rate associated with a negative experiences would likely result in stress-related behaviors (e.g., pricked or flattened back ears, eye whites shown, tail swishing, snorting vocalization, trembling, sweating, etc.) and/or avoidance behaviors (e.g., leaning or stepping away, startle/flight response, etc.) towards the enrichment items. This is supported by another study which found that horses in novel object tests that showed stress-related behaviors had significantly higher heart rate increase than horses that did not show stress behaviors (Safryghin et al., 2019). In our study, we noticed slight startle behavior when mirrors were first introduced, but did not notice any other stress or avoidance responses during enrichment treatments (personal observation). This is evidenced by our results which show horses spent 13.36% of the observations engaged with enrichment and frustration behavior significantly decreased from control. It could be argued that horses may still investigate enrichment if they are anxious because they want to assess the risk factor associated with the novel object. However, if this were the case, we would expect the interaction to be most prevalent in the morning when the item is first introduced and decrease as the day goes on because if the risk factor was deemed high after initial assessment, an anxious horse would most likely avoid the item in later exposures. Yet, horses interacted with the hay feeder and activity ball similarly across all times; the only significant difference was within the mirror treatment in which it was lower in the morning than during noon and afternoon. Furthermore, state-anxiety (high arousal, low valence) is often accompanied by an increase in respiration rate (Gomez et al., 2004), but our results found no difference in respiration. Limited research has been conducted analyzing physiological effects of enrichment items on horses; however, music as a form of enrichment has been shown to have positive effects on horse emotional states (Stachurska et al., 2015). Though, whereas we found evidence for increased heart rate, Kędzierski et al. (2017a, 2017b) found that race horses had lower heart rate, heart rate variability, and salivary cortisol concentrations when they were exposed to new age music, indicating a more relaxed state in the horses. Neveux et al. (2016) also saw a reduction in horse stress responses when exposed to transport and farriery while classical music was playing. While music seems to have the opposite arousal effect compared to the enrichment items in our study, these studies similarly suggest that certain types of species-specific enrichment could promote positively valenced affective states. Further research involving cognitive bias testing is needed to more accurately infer horse affective states in response to enrichment items; however, behavioral results of our study suggest that horses did not respond negatively to the tested enrichment items. Additionally, this discussion has overviewed several instances in which the items allowed horses to perform more species-typical behaviors, and therefore we conclude that the hay feeder, activity ball, and mirror can be evaluated as effective enrichment options.

While there were evident effects on behavior between enrichment treatments and control, there were limited differences between the three enrichment items. This suggests that the behaviors for which there were no significant differences between enrichment treatments at any time point (i.e., enrichment interaction, standing rest, social behavior, frustration behaviors, and other behaviors) are not affected by type of enrichment used, and all three items may be equally beneficial for encouraging changes in those behaviors. This is good news for facilities that may be on a tighter budget and unable to obtain "higher quality" or a larger variety of enrichment items, as it allows them to not feel restricted based on what they can provide while still feeling optimistic that they are providing some extent of improved welfare. There were, however, some significant differences between enrichment items for foraging, locomotion, and standing alert. Specifically, the hay feeder had significantly higher foraging than the activity ball and mirror at noon; the hay feeder and activity ball had significantly higher locomotion than the mirror in the morning, noon, and afternoon; and the hay feeder had significantly lower standing alert behavior than the activity ball but not the mirror in the afternoon. This indicates an enrichment type effect for these behaviors and facilities are recommended to use specific enrichment items to target behaviors and achieve their specific welfare goals based on individual horse needs. Taylor et al. (2022) also highlights this concept in their paper proposing an enrichment framework based on a literature review and industry survey in which they conclude that focusing enrichment strategies on specific animal welfare outcomes, rather than the intent, will assist owners in identifying the optimal enrichment for their purposes. Overall, the hay feeder appears to provide the most prominent changes in heart rate and behavior from control among the enrichment items at all time points and has a more species-typical activity budget that is similarly consistent throughout the day (Figure 4). This suggests that the hay feeder may be more effective as an enrichment item in fulfilling overall behavioral needs. However, throughout the discussion we have provided strong supporting evidence that the activity ball and mirror are also beneficial in promoting positive behavioral changes for improved welfare, and therefore we also recommend their use as effective enrichment.

7. Other Considerations and Study Limitations

During one of the research sessions, a horse had a vet visit checkup. While this single observation (hay feeder at noon for one horse) was excluded from the statistical analysis, it is possible the vet visit had lingering effects on the horse. The horse did not appear to have adverse side effects or behave differently than usual during the rest of the day (personal observation), so it is assumed that the event had minimal to no effects on subsequent research sessions for the horse that day; however, it is possible that effects existed and impacted those session results. Therefore, we note this for consideration as a study limitation that is recommended to be avoided in future research studies.

Laying down behavior was not recorded for any treatment and could not be assessed other than to say that the tested enrichment items did not influence this behavior. Additionally, there were no significant differences in defecation behavior for any analysis. Increased defecation can sometimes occur in response to stress or fear such as when horses are put in new environments or encounter a novel object (Momozawa et al., 2003). Thus, our study indicates that the hay feeder, activity ball, and mirror enrichment items do not cause these negative reactions in horses. There was also limited change in other behaviors (e.g., drinking, urinating, grooming, etc.), with the only significant difference being the hay feeder, activity ball, and mirror had significantly more of these behaviors in the afternoon compared to control and no difference amongst the three enrichments. The reason for this could be due to the increased foraging and locomotion, causing the horses to be thirstier at this time, inducing more drinking and subsequently urination. Regardless, the limited change in other behaviors indicates that the enrichment items did not have substantial influence on “side effects.”

In addition to the physiological and behavioral effects discussed, these enrichment items could be encouraging mental stimulation as well. The horse may need to problem solve and practice how to effectively grab the hay out of the feeder by manipulating its lips, all the while managing the feeder rolling around the stall and shifting where the hay openings are presented. It was observed that many of the interactions with the activity ball involved pushing it out of the way of their foraging, adding complexity to the behavior by providing opportunity for the horse to work and problem solve to obtain the food. Additionally, one horse attempted and was successful at picking the ball up and shook it around. Needing to manipulate its teeth and lips in a way to successfully pick up the ball presents a unique mental stimulation. Meanwhile, unlike the hay feeder or ball, the mirror provides stimulation through its factor of a novel concept (reflection) that the horse may not have previously been exposed to or encounter very often. As mentioned in the literature review, it is likely that horses can perceive mirror reflections due to the contrast of the images it is reflecting, particularly when there is movement via the horse or surroundings as this is something to which horses’ eyes are extremely sensitive (McGreevy, 2004). With this in mind, the mirror’s reflection may encourage the horse to investigate and try to figure out what it is and how the horse should emotionally and behaviorally react. While the hay feeder and activity ball had stronger overall behavioral effects due to those items increasing foraging and locomotion more than the mirror, the mirror may still be similarly effective as an enrichment item given its possible benefits of mental stimulation. From their study in which mirrors significantly reduced stereotypic weaving, McAfee et al. (2002) concluded that mirrors may alter the horse’s perception of the environment by creating the illusion of visual contact with conspecifics and thus influence the resultant responses to that environment. Additionally, Decker et al. (2023) who reviewed literature to assess enrichment effects when there is no active interaction concluded that active interaction does not always indicate welfare value and that items that elicit little to no interaction still have welfare benefits such as reducing stress. This further speaks to the value of enrichment as a form of mental stimulation, especially mirrors which may have less physical interaction than other enrichment items. Though, more research needs to be done to assess cognitive effects of enrichment.

For ease while conducting research sessions in which enrichment items were removed between sessions, we attached mirrors to stalls using hooks and rope. This is not recommended for official use as it allowed easy access for horses to grab the mirror with their lips and slide it along the bar and/or break off pieces of the mirror. While the horses never attempted to eat the mirror nor did the mirror cause a safety concern of injuring the horse (broken pieces were immediately taken out of the stall and any sharp edges of the in-tact mirror had padded gauze applied to them), this could be a concern if the horse is not being constantly monitored. Therefore, for official use we recommend securely attaching the mirror flush with a wall and attaching with double sided tape or screws. Alternately, managers could hang more natural or sturdy enrichments like the hay feeder or ball on a rope to encourage the similar tactile manipulations we unintentionally achieved with the mirror. More research would need to be done to determine how effects of this hanging type of enrichment would compare to the items tested in this study.

It is important to note that not all horses may respond the same way to an enrichment item. In the current study, one horse was observed to pick up the activity ball and shake it around. For budget purposes, we did not buy covers for the activity balls, but there are some you can buy with handles that would offer the option for the horse to pick up the ball more easily. As we only saw this from one horse, the appeal to pick up the ball may be personality dependent. Reactions to enrichment may also be dependent on demographic factors. A study by Huo, Yaemklang, et al. (2021) found that reactions to different kinds of enrichment were significantly different between horses that were known to perform stereotypic behaviors and those that did not show any abnormal behavior. Bulens et al. (2013) also found that younger horses and stallions showed significantly higher frequency of item related behavior towards sand filled bottles and rope, indicating that behavior may be associated with age and gender. Additionally, while the horses in the current study were a good representation of a typical barn management schedule, some facilities may operate differently which could alter results. Taking these points into consideration, managers should monitor horses when enrichment is introduced for the first time and should experiment to see what items works best for their horse.

Future studies should utilize larger sample sizes and continuous sampling methods to get a closer representation of daily activity budgets. Some topics needing further investigation have already been mentioned throughout the discussion; in addition to those, future studies could analyze other types of enrichment items to see how they compare; study breed, age, personality, and/or other equine species to see if and how effects change; and since we saw a reduction in frustration behaviors, a sample size of horses with stereotypic behaviors could be used to further analyze enrichment type effects on these behaviors.

8. Conclusion

Enrichment did not affect respiration rate, but did increase heart rate of stalled horses which indicates that horses were in a high arousal state when provided enrichment. The hay feeder showed the strongest physiological effects among the three enrichment items, and for all treatments, heart rate was not dependent on provision time except for the mirror having significantly lower heart rate in the evening compared to the rest of the day.

Overall, enrichment encouraged more species-typical activity budgets in stalled horses. All three enrichment items had positive effects compared to control and could be used as effective enrichment, though there were enrichment type effects for certain behaviors, so enrichment program strategies could be used to target and achieve specific management goals. However, the hay feeder appeared more effective at fulfilling overall behavioral needs. Behavioral effects of enrichment were present at all time points throughout the day except for the evening. However, there were more behavioral differences from control at noon and afternoon, so owners are recommended to prioritize giving enrichment when horses do not have daily meal provisions.

In conclusion, this study gives evidence that providing enrichment has a positive effect on physiology and behavior and is recommended to improve the behavioral needs, health, and overall welfare of stalled horses. Though it is important to consider that individual personalities and needs may shape how a horse interacts with an enrichment item, so owners should experiment to see which one works best for their horse. More research is recommended to study other enrichment items, affective states, and long-term effects.

9. References

Alligood, C., & Leighty, K. (2015). Putting the "E" in SPIDER: Evolving trends in the evaluation of environmental enrichment efficacy in zoological settings. *Animal Behavior and Cognition, 2*(3), 200–217. https://doi.org/10.12966/abc.08.01.2015

American Wild Horse Conservation. (2017). *Frequently Asked Questions*. Retrieved September 9, 2023, from https://americanwildhorse.org/faq

Art, T., Desmecht, D., Amory, H., & Lekeux, P. (1990). Synchronization of locomotion and respiration in trotting ponies. *Journal of Veterinary Medicine, 37*(2), 97–103. https://doi.org/10.1111/j.1439-0442.1990.tb00880.x

Brambell, F. W. R. (1965). *Report of the Technical Committee of Enquiry into the Welfare of Livestock Kept under Intensive Conditions.* Her Majesty's Stationery Office. Retrieved January 2, 2024, from https://archive.org/details/b3217276x/mode/2up

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2014). Fitting linear mixed-effects models using lme4. *arXiv preprint arXiv:1406.5823*. https://doi.org/10.48550/arXiv.1406.5823

Baumgartner, M., Boisson, T., Erhard, M. H., & Zeitler-Feicht, M. H. (2020). Common feeding practices pose a risk to the welfare of horses when kept on non-edible bedding. *Animals, 10*(3), 411. https://doi.org/10.3390/ani10030411

Benhajali, H., Richard-Yris, M. A., Ezzaouia, M., Charfi, F., & Hausberger, M. (2009). Foraging opportunity: A crucial criterion for horse welfare?. *Animal, 3*(9), 1308–1312. https://doi.org/10.1017/S1751731109004820

Boyd, L. E., Carbonaro, D. A., & Houpt, K. A. (1988). The 24-hour time budget of Przewalski horses. *Applied Animal Behaviour Science*, *21*(1–2), 5–17. https://doi.org/10.1016/0168-1591(88)90098-6

Browning, H. (2020). The natural behavior debate: Two conceptions of animal welfare. *Journal of Applied Animal Welfare Science*, *23*(3), 325–337. https://doi.org/10.1080/10888705.2019.1672552

Broom, D. M. (1991). Animal welfare: Concepts and measurement. *Journal of Animal Science*, *69*, 4167–4175. https://doi.org/10.2527/1991.69104167x

Bulens, A., Beirendonck, S. V., Thielen, J. V., & Driessen, B. (2013). The enriching effect of non-commercial items in stabled horses. *Applied Animal Behaviour Science, 143*, 46–51. http://doi.org/10.1016/j.applanim.2012.11.012

Bulens, A., Dams, A., Beirendonck, S. V., Thielen, J. V., & Driessen, B. (2015). A preliminary study on the long-term interest of horses in ropes and Jolly Balls. *Journal of Veterinary Behavior, 10*, 83–86. http://doi.org/10.1016/j.jveb.2014.08.003

Campbell-Thompson, M. L., & Merritt, A. M. (1987). Effect of ranitidine on gastric acid secretion in young male horses. *American Journal of Veterinary Research, 48*, 1511–1515. https://doi.org/10.1111/eve.12505

Camargo, F., Gombeski Jr, W. R., Barger, P., Jehlik, C., Wiemers, H., Mead, J., & Lawyer, A. (2018). Horse-related injuries: Causes, preventability, and where educational efforts should be focused. *Cogent Food & Agriculture, 4*(1), 1432168. https://doi.org/10.1080/23311932.2018.1432168

Clark, F. E. (2017). Cognitive enrichment and welfare: Current approaches and future directions. *Animal Behavior and Cognition, 4*(1), 52–71. https://doi.org/10.12966/abc.05.02.2017

Cooper, J. J., & McGreevy, P. D. (2002). Stereotypic behavior in the stabled horse: Cause, effects, and prevention without compromising horse welfare. In N. Waran (Ed.), *The welfare of horses* (pp. 99–124). Kluwer Academic Publishers.

Dawkins, M. S. (1998). Evolution and Animal Welfare. *The Quarterly Review of Biology*, *73*(3), 305–328. https://doi.org/10.1086/420307

Decker, S., Lavery, J. M., & Mason, G. J. (2023). Don't use it? Don't lose it! Why active use if not required for stimuli, resources or "enrichments" to have welfare value. *Zoo Biology, 42*(4), 467–475. https://doi.org/10.1002/zoo.21756

Ellis, A. D., Redgate, S., Zinchenko, S., Owen, H., Barfoot, C., & Harris, P. (2015). The effect of presenting forage in multi-layered haynets and at multiple sites on night time budgets of stabled horses. *Applied Animal Behaviour Science*, *171*, 108–116. http://doi.org/10.1016/j.applanim.2015.08.012

Endenburg, N. (1999). Perceptions and attitudes towards horses in European societies. *Equine Veterinary Journal, 28*, 38–41. https://doi.org/10.1111/j.2042-3306.1999.tb05154.x

Endler, N. S., Edwards, J. M., Vitelli, R., & Parker, J. D. A. (1989). Assessment of state and trait anxiety: Endler multidimensional anxiety scales. *Anxiety Research, 2*(1), 1–14. https://doi.org/10.1037/0022-3514.60.6.919

European Respiratory Society. (2016). Your lungs and exercise. *Breathe (Sheff), 12*, 97–100. https://doi.org/10.1183/20734735.ELF121

Franklin, S. H., Van Erck-Westergren, E., & Bayly, W. M. (2012). Respiratory responses to exercise in the horse. *Equine Veterinary Journal, 44*, 726–732. https://doi.org/10.1111/j.2042-3306.2012.00666.x

Fraser, D. (2008). Understanding animal welfare. *Acta Veterinaria Scandinavica, 50*(1), S1. https://doi.org/10.1186/1751-0147-50-S1-S1

Fraser, D., Weary, D. M., Pajor, E. A., & Milligan, B. N. (1997). A scientific conception of animal welfare that reflects ethical concerns. *Animal Welfare, 6*, 187–205. https://doi.org/10.1017/S0962728600019795

Friard, O. & Gamba, M. (2022). Behavioral Observation Research Interactive Software (BORIS; Version 7.13.6) [Computer software]. Retrieved from https://www.boris.unito.it/

Garner, J. P. (2005). Stereotypies and other abnormal repetitive behaviors: Potential impact on validity, reliability, and replicability of scientific outcomes. *Institute for Laboratory Animal Research Journal*, *46*(2), 106–117. https://doi.org/10.1093/ilar.46.2.106

Gellman, K., & Ruina, A. (2022). Standing horse posture: A longer stance is more stable. *Biology Open*, *11*(4), bio059139. https://doi.org/10.1242/bio.059139

Gomez, P., Stahel, W. A., & Danuser, B. (2004). Respiratory responses during affective picture viewing. *Biological Psychology, 67*, 359–373. https://doi.org/10.1016/s0301-0511(04)00074-2

Goodwin, D. (2002). Horse behaviour: evolution, domestication and feralisation. In N. Waran (Ed.), *The welfare of horses* (pp. 1–18). Kluwer Academic Publishers.

Goodwin, D., Davidson, H. P. B., & Harris, P. (2002). Foraging enrichment for stabled horses: Effects on behaviour and selection. *Equine Veterinary Journal*, *34*(7), 686–691. https://doi.org/10.2746/042516402776250450

Grime, E., Northrop, A., Rosbotham, M. E., & Howells, K. L. (2004). The effect of mirrors for the control of stereotypic weaving in the stabled horse. *BSAP Occasional Publication, 32,* 185–186. https://doi.org/10.1017/S0263967X00041410

Harris, P. A. (1999). How understanding the digestive process can help minimise digestive disturbances due to diet and feeding practices. In P. A. Harris, G. Gomarsall, H. P. B. Davidson, & R. Green (Eds.), *Proceedings of the BEVA specialist days on nutrition and behaviour* (pp. 45–50). Equine Veterinary Journal Ltd.

Hartman, N., & Greening, L. M. (2019). A Preliminary Study Investigating the Influence of Auditory Stimulation on the Occurrence of Nocturnal Equine Sleep-Related Behavior in Stabled Horses. *Journal of Equine Veterinary Science*, *82*, 102782. https://doi.org/10.1016/j.jevs.2019.07.003

Hemsworth, L. M., Jongman, E. C., & Coleman, G. J. (2021). The human-horse relationship: Identifying the antecedents of horse owner attitudes towards horse husbandry and management behaviour. *Animals, 11*, 278. https://doi.org/10.3390/ani11020278

Henderson, J. V. (2007). Don't fence me in: Managing psychological well-being for elite performance horses. *Journal of Applied Animal Welfare Science, 10*(4), 309–329. https://doi.org/10.1080/10888700701555576

Henderson, J. V., & Waran, N. K. (2001). Reducing equine stereotypies using an Equiball™. *Animal Welfare, 10,* 73–80. https://doi.org/10.1017/S0962728600023277

Henneke, D. R., Potter, G. D., Kreider, J. L., & Yeates, B. F. (1983). Relationship between condition score, physical measurements and body fat percentage in mares. *Equine Veterinary Journal, 15*(4), 371–372. https://doi.org/10.1111/j.2042-3306.1983.tb01826.x

Horback, K. M., & Parsons, T. D. (2019). Judgement bias testing in group-housed gestating sows. *Behavioural Processes, 159,* 86–92. https://doi.org/10.1016/j.beproc.2018.12.021

Hotchkiss, J. W., Reid, S. W. J., & Christley, R. M. (2007). A survey of horse owners in Great Britain regarding horses in their care. Part 1: Horse demographic characteristics and management. *Equine Veterinary Journal, 39*(4), 294–300. https://doi.org/10.2746/042516407x177538

Hothorn, T., Bretz, F., & Westfall, P. (2008). Simultaneous inference in general parametric models. *Biometrical Journal: Journal of Mathematical Methods in Biosciences, 50*(3), 346–363. https://doi.org/10.1002/bimj.200810425

Houpt, K., Marrow, M., & Seeliger, M. (2000). A preliminary study of the effect of music on equine behavior. *Journal of Equine Veterinary Science*, *20*(11), 691–737. https://doi.org/10.1016/S0737-0806(00)80155-0

Hox, J. J., Moerbeek, M., Kluytmans, A., & van de Schoot, R. (2014). Analyzing indirect effects in cluster randomized trials. The effect of estimation method, number of groups and group sizes on accuracy and power. *Frontiers in Psychology, 5*, 78. https://doi.org/10.3389/fpsyg.2014.00078

Huo, X., Wongkwanklom, M., Phonraksa, T., & Na-Lampang, P. (2021). Effects of playing classical music on behavior of stabled horses. *Veterinary Integrative Sciences*, *19*(2), 259–267. https://doi.org/10.12982/VIS.2021.023

Huo, X., Yaemklang, S., Pimmai, P., Kupittayanant, P., & Na-Lampang, P. (2021). A preliminary study of the effects of enrichment on stereotypic and non-stereotypic horses. *Veterinary Integrative Sciences, 19*(3), 581–590. https://doi.org/10.12982/VIS.2021.045

Jørgensen, G. H. M., & Bøe, K. E. (2007a). A note on the effect of daily exercise and paddock size on the behaviour of domestic horses (*Equus caballus*). *Applied Animal Behaviour Science, 107*, 166–173. https://doi.org/10.1016/j.applanim.2006.09.025

Jørgensen, G. H. M., & Bøe, K. E. (2007b). Individual paddocks versus social enclosure for horses. *EAAP Scientific Series, 122*, 79–83. https://doi.org/10.3920/9789086866144_010

Jørgensen, G. H. M., Liestøl, S. H. O., & Bøe, K. E. (2010). Use of different items of "enrichment" for individual and group kept horses. *Journal of Veterinary Behavior, 5*(4), 216. https://doi.org/10.1016/j.jveb.2009.10.039

Jørgensen, G. H. M., Liestøl, S. H. O., & Bøe, K. E. (2011). Effects of enrichment items on activity and social interactions in domestic horses (*Equus caballus*). *Applied Animal Behaviour Science*, *129*(2–4), 100–110. https://doi.org/10.1016/j.applanim.2010.11.004

Karaer, M. C., Čebulj-Kadunc, N., & Snoj, T. (2023). Stress in wildlife: comparison of the stress response among domestic, captive, and free-ranging animals. *Frontiers in Veterinary Science, 10,* 1167016. https://doi.org/10.3389/fvets.2023.1167016

Kędzierski, W., Janczarek, I., Stachurska, A., & Wilk, I. (2017a). Comparison of effects of different relaxing massage frequencies and different music hours on reducing stress level in race horses. *Journal of Equine Veterinary Science*, *53*, 100–107. http://doi.org/10.1016/j.jevs.2017.02.004

Kędzierski, W., Janczarek, I., Stachurska, A., & Wilk, I. (2017b). Massage or music meant to be relaxing, result in lowering salivary cortisol concentration in race horses. *Pferdeheilkunde Equine Medicine*, *33*(2), 146–151. https://doi.org/10.21836/PEM20170206

Kenney, T. (2018). *The effects of enrichment items on cribbing in Equus caballus*. [Honors thesis, Alfred University]. AURA: Alfred University Research & Archives. http://hdl.handle.net/10829/8208

Knubben, J. M., Fürst, A., Gygax, L., & Stauffacher, M. (2008). Bite and kick injuries in horses: Prevalence, risk factors and prevention. *Equine Veterinary Journal*, *40*(3), 219–223. https://doi.org/10.2746/042516408x253118

Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics, 33*, 159–174. https://doi.org/10.2307/2529310

Lee, S. Y., & Song, X. Y. (2004). Evaluation of the Bayesian and maximum likelihood approaches in analyzing structural equation models with small sample sizes. *Multivariate Behavioral Research, 39*(4), 653–686. https://doi.org/10.1207/s15327906mbr3904_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

Lewis, H. E. (2021, June 25). *The basics of horse stall design*. Stable Management. Retrieved October 24, 2023, from https://stablemanagement.com/barns-grounds/the-basics-of-horse-stall-design/

Martinson, K., Hathaway, M., Wilson, J. H., Gilkerson, B., Peterson, P. R., & Del Vecchio, R. (2006). University of Minnesota horse owner survey: Building an equine extension program. *Journal of Extension, 44*(6), 1–8. https://archives.joe.org/joe/2006december/rb4.php

McAfee, L. M., Mills, D. S., & Cooper, J. J. (2002). The use of mirrors for the control of stereotypic weaving behaviour in the stabled horse. *Applied Animal Behaviour Science*, *78*(2–4), 159–173. https://doi.org/10.1016/S0168-1591(02)00086-2

McBride, S. D., & Hemmings, A. (2009). A neurologic perspective of equine stereotypy. *Journal of Equine Veterinary Science, 29*, 10–16. https://doi.org/10.1016/j.jevs.2008.11.008

McClure, S. R. (2016). *Equine gastric ulcers: Special care and nutrition*. American Association of Equine Practitioners. Retrieved October 23, 2023, from https://aaep.org/horsehealth/equine-gastric-ulcers-special-care-and-nutrition

McGreevy, P. D., Cripps, P. J., French, N. P., Green, L. E., & Nicol, C. J. (1995). Management factors associated with stereotypic and redirected behaviour in the thoroughbred horse. *Equine Veterinary Journal, 27*, 86–91. https://doi.org/10.1111/J.2042-3306.1995.TB03041.X

McGreevy, P. (2004). *Equine behavior: A guide for veterinarians and equine scientists*. Saunders Ltd.

McPhee, M. E., & Carlstead, K. (2010). The importance of maintaining natural behaviors in captive mammals. In D. Kleiman, K. Thompson, & K. Baer (Eds.), *Wild mammals in captivity: Principles and techniques for zoo management* (2nd ed., pp. 303–313). University of Chicago Press.

Mellen, J., & MacPhee, M. S. (2001). Philosophy of environmental enrichment: Past, present, and future. *Zoo Biology, 20*, 211–226. https://doi.org/10.1002/zoo.1021

Mills, D., & Davenport, K. (2002). The effect of a neighbouring conspecific *versus* the use of a mirror for the control of stereotypic weaving behaviour in the stabled horse. *Animal Science*, *74*(1), 95–101. https://doi.org/10.1017/S1357729800052255

Momozawa, Y., Ono, K., Sato, F., Kikusui, T., Takeuchi, Y., Mori, Y., et al. (2003). Assessment of equine temperament by a questionnaire survey to caretakers and evaluation of its reliability by simultaneous behavior test. *Applied Animal Behaviour Science*, *84*, 127–138. https://doi.org/10.1016/j.applanim.2003.08.001

Murray, M.J., & Schusser, G. (1993). Application of gastric pH-metry in horses: Measurement of 24 hour gastric pH in horses fed, fasted, and treated with ranitidine. *Journal of Veterinary Internal Medicine, 6*, 133. https://doi.org/10.1111/j.2042-3306.1993.tb02983.x

Neveux, C., Ferard, M., Dickel, L., Bouet, V., Petit, O., & Valenchon, M. (2016). Classical music reduces acute stress of domestic horses. *Journal of Veterinary Behavior*, *15*, 81. https://doi.org/10.1016/j.jveb.2016.08.019

Newberry, R. C. (1995). Environmental enrichment: Increasing the biological relevance of captive environments. *Applied Animal Behaviour Science, 44*(2–4), 229–243. https://doi.org/10.1016/0168-1591(95)00616-Z

Protequus. (2021). Nightwatch® Smart Halter™ [Equipment and software]. Protequus, Austin, Texas. https://smarthalter.com

Ogilvie-Graham, T. S. (1994). *Time budget studies in stalled horses.* [Doctoral dissertation, University of Edinburgh]. Edinburgh Research Archive. http://hdl.handle.net/1842/29924

Orsini, J. (2000). Gastric ulceration in the mature horse: A review. *Equine Veterinary Education*, *12*(1), 24–27. https://doi.org/10.1111/j.2042-3292.2000.tb01759.x

Park, S. K., Jung, H. J., Choi, Y. L., Kwon, O. S., Jung, Y. H., Cho, C., & Yoon, M. (2013). The effect of living conditions on stress and behavior of horses. *Journal of Animal Science and Technology, 55*(4), 325–330. https://doi.org/10.5187/JAST.2013.55.4.325

Perdue, B. M., Sherwen, S. L., & Maple, T. L. (2020). The science and practice of captive animal welfare [Editorial]. *Frontiers in Psychology, 11*, 1851. https://doi.org/10.3389/fpsyg.2020.01851

Posta, J., Komlósi, I., & Mihók, S. (2009). Breeding value estimation in the Hungarian sport horse population. *The Veterinary Journal, 181*, 19–23. https://doi.org/10.1016/j.tvjl.2009.03.006

Price, E. O. (1984). Behavioral aspects of animal domestication. *The Quarterly Review of Biology, 59*(1), 1–32. https://doi.org/10.1086/413673

R Core Team. (2013). R: A Language and Environment for Statistical Computing (Version 3.3.1) [Computer software]. R Foundation for Statistical Computing, Vienna, Austria. http://www.R-project.org/

Ransom, J. I., & Cade, B. S. (2009). Quantifying equid behavior: A research ethogram for free-roaming feral horses. *US Geological Survey Techniques and Methods 2-A9* (p. 23). United States Geological Survey. https://doi.org/10.3133/tm2A9

Rochais, C., Henry, S., & Hausberger, M. (2018). "Hay-bags" and "slow feeders": Testing their impact on horse behaviour and welfare. *Applied Animal Behaviour Science*, *198*, 52–59. http://doi.org/10.1016/j.applanim.2017.09.019

Ross, M., Proudfoot, K., Merkies, K., Elsohaby, I., Mills, M., Macmillan, K., Mckenna, S., & Ritter, C. (2023). Horse housing on Prince Edward Island, Canada: Attitudes and experiences related to keeping horses outdoors and in groups. *Animals, 13*(2), 275. https://doi.org/10.3390/ani13020275

Russell, J. A. (1980). A circumplex model of affect. *Journal of Personality and Social Psychology, 39*(6), 1161–1178. http://doi.org/10.1037/h0077714

Safryghin, A., Hebesberger, D. V., & Wascher, C. A. (2019). Testing for behavioral and physiological responses of domestic horses (*Equus caballus*) across different contexts – consistency over time and effects of context. *Frontiers in Psychology*, *10*, 849. https://doi.org/10.3389/fpsyg.2019.00849

Shepherdson, D. (1998). Tracing the path of environmental enrichment in zoos. In D. Shepherdson, J. D. Mellen, & M. Hutchins (Eds.), *Second nature: Environmental enrichment for captive animals* (pp. 1–13). Smithsonian Institution Press.

Søndergaard, E., Jensen, M. B., & Nicol, C. J. (2011). Motivation for social contact in horses measured by operant conditioning. *Applied Animal Behaviour Science, 132*, 131–137. https://doi.org/10.1016/j.applanim.2011.04.007

Stachurska, A., Janczarek, I., Wilk, I., & Kędzierski, W. (2015). Does music influence emotional state in race horses?. *Journal of Equine Veterinary Science, 35*, 650–656. http://doi.org/10.1016/j.jevs.2015.06.008

Tavernier, A. (1988). Advantages of BLUP animal model for breeding value estimation in horses. *Livestock Production Science, 20*, 149–160. https://doi.org/10.1016/0301-6226(88)90059-0

Taylor, P. S., Schrobback, P., Verdon, M., & Lee, C. (2022). An effective environmental enrichment framework for the continual improvement of production animal welfare. *Animal Welfare, 32*(e14), 1–11. https://doi.org/10.1017/awf.2023.5

Thorne, J., Goodwin, D., Kennedy, M., Davidson, H., & Harris, P. (2005). Foraging enrichment for individually housed horses: Practicality and effects on behaviour. *Applied Animal Behaviour Science*, *94*(1–2), 149–164. https://doi.org/10.1016/j.applanim.2005.02.002

Villas-Boas, J. D., Dias, D. P. M., Ignacio Trigo, P., dos Santos Almeida, N. A., de Almeida, F. Q., & de Medeiros, M. A. (2016). Behavioural, endocrine and cardiac autonomic responses to a model of startle in horses. *Applied Animal Behaviour Science, 174*, 76–82. https://doi.org/10.1016/j.applanim.2015.10.005

Villas-Boas, J. D., Dias, D. P. M., de Almeida, F. Q., & de Medeiros, M. A. (2022). Different equestrian sports horses' reactivity to startle. *Journal of Veterinary Behaviour, 57*, 16–23. https://doi.org/10.1016/j.jveb.2022.07.016

Waring, G. H. (2003). *Horse behavior*. Noyes Publications/William Andrew Publishing.

Webster, J. (2016). Animal welfare: Freedoms, dominions and "a life worth living". *Animals, 6*(6), 35. https://doi.org/10.3390/ani6060035

Wells, D. (2009). Sensory stimulation as environmental enrichment for captive animals: A review. *Applied Animal Behaviour Science, 118*, 1–11. https://doi.org/10.1016/j.applanim.2009.01.002

Werhahn, H., Hessel, E. F., & Van den Weghe, H. F. A. (2012). Competition horses housed in single stalls (II): Effects of free exercise on the behavior in the stable, the behavior during training, and the degree of stress. *Journal of Equine Veterinary Science, 32*(1), 22–31. https://doi.org/10.1016/j.jevs.2011.06.009

Wiepkema, P. R. (1987). Behavioural aspects of stress. In P. R. Wiepkema & P. W. M. van Adrichem (Eds.), *Biology of stress in farm animals: An integrative approach* (pp. 113–133). Martinus Nijhoff Publishers.

Winskill, L. C., Waran, N. K., & Young, R. J. (1996). The effect of a foraging device (a modified 'Edinburgh Foodball') on the behaviour of the stabled horse. *Applied Animal Behaviour Science, 48*, 25–35. https://doi.org/10.1016/0168-1591(95)01021-1

Wiśniewska, M., Janczarek, I., Wilk, I., & Wnuk-Pawlak, E. (2019). Use of music therapy in aiding the relaxation of geriatric horses. *Journal of Equine Veterinary Science*, *78*, 89–93. https://doi.org/10.1016/j.jevs.2018.12.011

World Weather. (2022, April). *Weather in Davis (California) in April 2022*. Retrieved July 23, 2023, from https://world-weather.info/forecast/usa/davis/april-2022/

Young, R. J. (2003). *Environmental enrichment for captive animals.* Blackwell Science Ltd.

Table 1. Test subjects' demographic information and group assignments for enrichment treatments.

Horse (n=9)	Group[a–c]	Sex	Age (years)
Bailey	A	Mare	5
Fargo	A	Gelding	7
Flash	A	Gelding	17
Bella	B	Mare	14
39	B	Gelding	14
Allen	B	Gelding	20
Rogue	C	Mare	6
Trevor	C	Gelding	7
Sparky	C	Gelding	12

[a] Group A (n=3) received the following enrichment treatment order: hay feeder (trial 1), mirror (trial 2), activity ball (trial 3). For each trial, day 1 = control, day 2 = enrichment.
[b] Group B (n=3) received the following enrichment treatment order: mirror (trial 1), activity ball (trial 2), hay feeder (trial 3). For each trial, day 1 = control, day 2 = enrichment.
[c] Group C (n=3) received the following enrichment treatment order: activity ball (trial 1), hay feeder (trial 2), mirror (trial 3). For each trial, day 1 = control, day 2 = enrichment.

Table 2. Enrichment items used for the hay feeder, activity ball, and mirror treatments.

Item		**Color**	**Size**
	High Country Plastics Hay Play™ slow feeding forage ball	red	16 in. diameter
	Horsemen's Pride mega ball	red	25 in. diameter
	Marketing Holders acrylic mirror sheet	n/a	24 in. x 24 in.

Table 3. Ethogram of behaviors recorded during observations (adapted from McGreevy, 2004; Rochais et al., 2018; Waring, 2003).

Behavior	Description
Foraging	Horse is searching for and/or grazing/chewing/eating food that did not directly come from the hay feeder enrichment item; Horse may be moving[a] around the stall with head below the withers and muzzle close to the ground or shifting through the substrate, may or may not be actively chewing; Horse may be standing stationary[b] with head either below or above the withers and is chewing food
Locomotion	Horse has taken at least two steps forward or backward and is walking or running around the stall without interacting with the enrichment item
Standing Alert	Horse is standing stationary with head and neck erect above the withers; ears are usually upright and forward facing, but one or both may be turned sideways or backward facing depending on sound direction and the horse's focus of attention
Standing Rest	Horse is standing stationary with head drooped equal to or below the withers; one hind leg may be relaxed with hoof resting on its toe; eyes may be open, half closed, or fully closed
Lying Down	Horse is lying on the ground partially on its side with legs tucked underneath and head may be lowered or upright; Horse is lying flat on its side with legs stretched straight and head is flat to the ground; Eyes may be open, half closed, or fully closed
Defecation	Horse is excreting feces
Social Behavior	Horse positively or negatively interacts with conspecifics or humans in one of the following ways: • Touch: Muzzle physically touches any body part of recipient, but there is no manipulation of the lips • Grooming: Muzzle physically touches any body part of a recipient conspecific, involves lip and mouth manipulation in which the horse brushes, licks, bites, and/or pulls the recipient's skin or hair; Recipient may or may not groom the initiator simultaneously (allogrooming) • Charge: Horse moves briskly in recipient's direction but does not make physical contact • Bite/Nip: Horse quickly moves head towards recipient and bites with lips or teeth onto any body part of the recipient; Conspecifics may or may not bite back, which may start a back-and-forth biting bout between the two horses

Table 3 (continued).

Behavior	Description
Enrichment Interaction	Horse is interacting with enrichment item in one of the following ways: • Investigation: Horse is standing still and looking at the enrichment item; horse sniffs item with its muzzle close to or touching the enrichment item; and/or horse startles (denoted by muscles tensing and jolt of the body and/or head with ears forward facing) and briskly retreats from the enrichment item • Push: Horse pushes enrichment item with its muzzle, leg, or body • Rub: Horse rubs its head, leg, or body on the enrichment item • Lick: Horse licks the enrichment item • Bite: Horse is biting on the enrichment item • Forage: Horse is searching for and/or grazing/chewing/eating food directly from the enrichment item[c]
Frustration Behavior	Horse expresses frustration in one of the following ways: • Stomping: Horse abruptly stomps foot on the ground, may be repetitive • Kicking: Horse is kicking, may or may not be directed towards a conspecific or person • Biting: Horse bites an object that is not food or enrichment item, may be repetitive • Pawing: Horse paws at the ground with its hoof, not followed by rolling • Repetitive Behavior: Horse expresses repetitive behavior in the form of yawning, licking, or stereotypic behavior
Other	Any behavior not otherwise listed (e.g., drinking, urinating, grooming, rolling)
Out of view	Horse's head is lowered and completely hidden from view behind the stall wall and behavior is unable to be accurately recorded

[a] If the horse was in locomotion while foraging, the behavior was scored as foraging if the horse's head remained below the withers and was scored as locomotion if the horse's head raised above the withers at the time of the scan.

[b] If the horse was standing alert while foraging, the behavior was scored as foraging if the horse was actively chewing and was scored as standing alert if the horse was not actively chewing at the time of the scan.

[c] As it was difficult to identify if hay on the ground came from the enrichment or daily provision, horses foraging from the ground was scored as regular foraging behavior, not enrichment interaction.

Table 4: Heart rate (mean beats per minute ± SE) of stalled horses when provided a hay feeder, activity ball, mirror, and control during 30-minute observations at different times of the day. P-values represent Tukey HSD post-hoc multi-comparisons.

	Hay Feeder (HF)	**Activity Ball (AB)**	**Mirror (MIR)**	**Control (CON)**	**p-value (HF v. AB)**	**p-value (HF v. MIR)**	**p-value (HF v. CON)**	**p-value (AB v. MIR)**	**p-value (AB v. CON)**	**p-value (MIR v. CON)**
Morning	70.58 ± 3.52 ^{Ax}	53.58 ± 3.22 ^{Bx}	64.25 ± 5.06 ^{Bx}	29.52 ± 3.88 ^{Cx}	0.028	0.028	0.027	NS	0.027	0.022
Noon	71.96 ± 4.25 ^{Ax}	62.52 ± 4.25 ^{Ax}	55.85 ± 6.52 ^{Bx}	35.56 ± 4.58 ^{Cx}	NS	0.036	0.022	0.029	0.028	0.019
Afternoon	65.58 ± 6.52 ^{Ax}	54.22 ± 4.20 ^{Bx}	53.41 ± 4.58 ^{Bx}	37.89 ± 4.02 ^{Cx}	0.032	0.032	0.036	NS	0.027	0.024
Evening	75.85 ± 3.25 ^{Ax}	46.52 ± 3.85 ^{Bx}	42.52 ± 5.52 ^{By}	30.22 ± 3.99 ^{Cx}	0.039	0.041	0.038	NS	0.031	0.036

A-C different superscripts represent statistically significant difference within the same time of day [row] and across different treatments [column] at $p<0.05$. x-y different superscripts represent statistically significant difference within the same treatment [column] and across different times of the day [rows] at $p<0.05$.

Table 5: Respiration rate (mean breaths per minute ± SE) of stalled horses when provided a hay feeder, activity ball, mirror, and control during 30-minute observations at different times of the day. P-values represent Tukey HSD post-hoc multi-comparisons.

	Hay Feeder (HF)	**Activity Ball (AB)**	**Mirror (MIR)**	**Control (CON)**	**p-value (HF v. AB)**	**p-value (HF v. MIR)**	**p-value (HF v. CON)**	**p-value (AB v. MIR)**	**p-value (AB v. CON)**	**p-value (MIR v. CON)**
Morning	24.25 ± 3.22 ^{Ax}	20.52 ± 4.69 ^{Ax}	23.58 ± 5.63 ^{Ax}	22.58 ± 6.96 ^{Ax}	NS	NS	NS	NS	NS	NS
Noon	22.52 ± 2.96 ^{Ax}	22.56 ± 4.25 ^{Ax}	20.54 ± 4.52 ^{Ax}	19.88 ± 4.56 ^{Ax}	NS	NS	NS	NS	NS	NS
Afternoon	20.36 ± 4.01 ^{Ax}	18.63 ± 3.55 ^{Ax}	16.85 ± 3.69 ^{Ax}	17.85 ± 6.32 ^{Ax}	NS	NS	NS	NS	NS	NS
Evening	23.58 ± 3.55 ^{Ax}	16.22 ± 4.25 ^{Ax}	17.85 ± 6.63 ^{Ax}	23.85 ± 4.55 ^{Ax}	NS	NS	NS	NS	NS	NS

A-C different superscripts represent statistically significant difference within the same time of day [row] and across different treatments [column] at $p<0.05$. x-y different superscripts represent statistically significant difference within the same treatment [column] and across different times of the day [rows] at $p<0.05$.

Table 6. Percentage of observation (mean ± SE) the stalled horses spent performing each behavior when provided the hay feeder, activity ball, mirror, and control during 30-minute observations at different times of the day. P-values represent Tukey HSD post-hoc multi-comparisons.

	Hay Feeder (HF)	**Activity Ball (AB)**	**Mirror (MIR)**	**Control (CON)**	**p-value (HF v. AB)**	**p-value (HF v. MIR)**	**p-value (HF v. CON)**	**p-value (AB v. MIR)**	**p-value (AB v. CON)**	**p-value (MIR v. CON)**
Foraging:										
Morning	60.23 ± 5.25 ^{Ax}	53.85 ± 3.89 ^{Ax}	55.63 ± 6.03 ^{Ax}	49.85 ± 5.63 ^{Bx}	NS	NS	0.035	NS	0.038	0.041
Noon	57.99 ± 4.89 ^{Ax}	44.52 ± 4.25 ^{Bx}	45.66 ± 4.25 ^{Bx}	30.56 ± 4.58 ^{Cy}	0.036	0.029	0.029	NS	0.036	0.032
Afternoon	54.36 ± 3.55 ^{Ax}	45.23 ± 5.12 ^{Ax}	47.96 ± 7.52 ^{Ax}	31.58 ± 6.99 ^{By}	NS	NS	0.025	NS	0.023	0.039
Evening	59.85 ± 5.63 ^{Ax}	52.36 ± 6.63 ^{Ax}	53.59 ± 2.36 ^{Ax}	49.85 ± 6.87 ^{Ax}	NS	NS	NS	NS	NS	NS
Locomotion:										
Morning	7.23 ± 1.03 ^{Ay}	8.69 ± 1.56 ^{Ay}	4.25 ± 1.01 ^{Bx}	3.63 ± 1.03 ^{Bx}	NS	0.039	0.031	0.043	0.019	NS
Noon	14.03 ± 2.96 ^{Ax}	16.25 ± 2.52 ^{Ax}	3.69 ± 0.96 ^{Bx}	1.66 ± 0.69 ^{Bx}	NS	0.026	0.042	0.023	0.031	NS
Afternoon	15.99 ± 3.66 ^{Ax}	14.52 ± 3.33 ^{Ax}	4.03 ± 1.12 ^{Bx}	1.25 ± 0.23 ^{Bx}	NS	0.032	0.025	0.037	0.036	NS
Evening	5.66 ± 1.03 ^{Ay}	7.58 ± 1.01 ^{Ay}	9.85 ± 1.33 ^{Ax}	5.63 ± 1.03 ^{Ax}	NS	NS	NS	NS	NS	NS

Table 6 (continued).

	Hay Feeder (HF)	Activity Ball (AB)	Mirror (MIR)	Control (CON)	p-value (HF v. AB)	p-value (HF v. MIR)	p-value (HF v. CON)	p-value (AB v. MIR)	p-value (AB v. CON)	p-value (MIR v. CON)
Standing Alert:										
Morning	6.36 ± 1.22 ^{Ax}	10.03 ± 1.30 ^{Ax}	10.69 ± 1.58 ^{Ax}	14.66 ± 3.63 ^{Ax}	NS	NS	NS	NS	NS	NS
Noon	2.96 ± 0.36 ^{Bx}	9.68 ± 1.99 ^{Bx}	8.52 ± 1.03 ^{Bx}	25.63 ± 4.25 ^{Ax}	NS	NS	0.036	NS	0.032	0.037
Afternoon	3.96 ± 1.03 ^{Cx}	11.52 ± 2.30 ^{Bx}	7.85 ± 1.12 ^{BCx}	21.63 ± 2.33 ^{Ax}	0.036	NS	0.029	NS	0.028	0.022
Evening	4.36 ± 1.33 ^{Ax}	8.63 ± 1.36 ^{Ax}	11.25 ± 2.03 ^{Ax}	16.85 ± 1.63 ^{Ax}	NS	NS	NS	NS	NS	NS
Standing Rest:										
Morning	4.99 ± 1.12 ^{Bx}	6.55 ± 1.96 ^{Bx}	10.03 ± 2.03 ^{Bx}	16.99 ± 1.55 ^{Ax}	NS	NS	0.036	NS	0.035	0.042
Noon	2.03 ± 0.99 ^{Bx}	5.23 ± 1.55 ^{Bx}	8.52 ± 1.15 ^{Bx}	23.20 ± 2.03 ^{Ax}	NS	NS	0.036	NS	0.032	0.041
Afternoon	3.69 ± 1.30 ^{Bx}	4.25 ± 0.87 ^{Bx}	8.23 ± 1.12 ^{Bx}	21.99 ± 2.06 ^{Ax}	NS	NS	0.023	NS	0.026	0.023
Evening	4.63 ± 1.33 ^{Ax}	9.85 ± 1.58 ^{Ax}	11.36 ± 1.05 ^{Ax}	13.52 ± 1.85 ^{Ax}	NS	NS	NS	NS	NS	NS

Table 6 (continued).

	Hay Feeder (HF)	Activity Ball (AB)	Mirror (MIR)	Control (CON)	p-value (HF v. AB)	p-value (HF v. MIR)	p-value (HF v. CON)	p-value (AB v. MIR)	p-value (AB v. CON)	p-value (MIR v. CON)
Lying Down:										
Morning	0.00 ± 0.00 ^{Ax}	0.00 ± 0.00 ^{Ax}	0.00 ± 0.00 ^{Ax}	0.00 ± 0.00 ^{Ax}	NS	NS	NS	NS	NS	NS
Noon	0.00 ± 0.00 ^{Ax}	0.00 ± 0.00 ^{Ax}	0.00 ± 0.00 ^{Ax}	0.00 ± 0.00 ^{Ax}	NS	NS	NS	NS	NS	NS
Afternoon	0.00 ± 0.00 ^{Ax}	0.00 ± 0.00 ^{Ax}	0.00 ± 0.00 ^{Ax}	0.00 ± 0.00 ^{Ax}	NS	NS	NS	NS	NS	NS
Evening	0.00 ± 0.00 ^{Ax}	0.00 ± 0.00 ^{Ax}	0.00 ± 0.00 ^{Ax}	0.00 ± 0.00 ^{Ax}	NS	NS	NS	NS	NS	NS
Defecation:										
Morning	0.19 ± 0.03 ^{Ax}	0.19 ± 0.02 ^{Ax}	0.22 ± 0.01 ^{Ax}	0.16 ± 0.05 ^{Ax}	NS	NS	NS	NS	NS	NS
Noon	0.25 ± 0.02 ^{Ax}	0.25 ± 0.02 ^{Ax}	0.24 ± 0.03 ^{Ax}	0.21 ± 0.06 ^{Ax}	NS	NS	NS	NS	NS	NS
Afternoon	0.17 ± 0.06 ^{Ax}	0.22 ± 0.03 ^{Ax}	0.19 ± 0.06 ^{Ax}	0.15 ± 0.03 ^{Ax}	NS	NS	NS	NS	NS	NS
Evening	0.16 ± 0.04 ^{Ax}	0.13 ± 0.06 ^{Ax}	0.17 ± 0.09 ^{Ax}	0.23 ± 0.04 ^{Ax}	NS	NS	NS	NS	NS	NS

Table 6 (continued).

	Hay Feeder (HF)	Activity Ball (AB)	Mirror (MIR)	Control (CON)	p-value (HF v. AB)	p-value (HF v. MIR)	p-value (HF v. CON)	p-value (AB v. MIR)	p-value (AB v. CON)	p-value (MIR v. CON)
Social Behavior:										
Morning	3.23 ± 0.55 ^{Ax}	2.63 ± 0.88 ^{Ax}	5.03 ± 1.01 ^{Ax}	4.99 ± 1.03 ^{Ax}	NS	NS	NS	NS	NS	NS
Noon	0.41 ± 0.03 ^{Bx}	2.25 ± 0.96 ^{Bx}	1.13 ± 0.65 ^{Bx}	8.25 ± 1.25 ^{Ax}	NS	NS	0.032	NS	0.025	0.039
Afternoon	0.56 ± 0.06 ^{Bx}	1.12 ± 0.13 ^{Bx}	3.66 ± 1.03 ^{Bx}	9.12 ± 1.63 ^{Ax}	NS	NS	0.021	NS	0.029	0.023
Evening	1.33 ± 0.12 ^{Ax}	3.03 ± 0.23 ^{Ax}	3.03 ± 0.96 ^{Ax}	3.52 ± 0.85 ^{Ax}	NS	NS	NS	NS	NS	NS
Enrichment Interaction:										
Morning	11.23 ± 1.23 ^{Ax}	10.63 ± 2.03 ^{Ax}	7.85 ± 1.36 ^{Ay}	0.00 ± 0.00 ^{Bx}	NS	NS	N/A	NS	N/A	N/A
Noon	16.58 ± 1.99 ^{Ax}	14.25 ± 3.25 ^{Ax}	19.52 ± 4.52 ^{Ax}	0.00 ± 0.00 ^{Bx}	NS	NS	N/A	NS	N/A	N/A
Afternoon	15.66 ± 2.02 ^{Ax}	13.25 ± 2.66 ^{Ax}	18.52 ± 3.63 ^{Ax}	0.00 ± 0.00 ^{Bx}	NS	NS	N/A	NS	N/A	N/A
Evening	16.99 ± 1.58 ^{Ax}	10.63 ± 3.69 ^{Ax}	5.22 ± 1.03 ^{Ay}	0.00 ± 0.00 ^{Bx}	NS	NS	N/A	NS	N/A	N/A

Table 6 (continued).

	Hay Feeder (HF)	**Activity Ball (AB)**	**Mirror (MIR)**	**Control (CON)**	**p-value (HF v. AB)**	**p-value (HF v. MIR)**	**p-value (HF v. CON)**	**p-value (AB v. MIR)**	**p-value (AB v. CON)**	**p-value (MIR v. CON)**
Frustration Behavior:										
Morning	2.66 ± 0.23 ^{Ax}	3.56 ± 0.22 ^{Ax}	4.36 ± 1.01 ^{Ax}	4.52 ± 1.12 ^{Ax}	NS	NS	NS	NS	NS	NS
Noon	0.66 ± 0.06 ^{Bx}	1.03 ± 0.12 ^{Bx}	2.99 ± 0.63 ^{Bx}	7.89 ± 1.13 ^{Ax}	NS	NS	0.035	NS	0.041	0.027
Afternoon	0.89 ± 0.09 ^{Bx}	1.12 ± 0.13 ^{Bx}	2.03 ± 0.16 ^{Bx}	11.96 ± 1.69 ^{Ax}	NS	NS	0.031	NS	0.041	0.043
Evening	2.01 ± 0.63 ^{Ax}	3.03 ± 0.96 ^{Ax}	2.69 ± 0.25 ^{Ax}	3.96 ± 0.98 ^{Ax}	NS	NS	NS	NS	NS	NS
Other Behavior:										
Morning	3.88 ± 1.81 ^{Ax}	3.87 ± 0.85 ^{Ax}	1.94 ± 0.68 ^{Ax}	5.20 ± 1.22 ^{Ax}	NS	NS	NS	NS	NS	NS
Noon	5.09 ± 1.66 ^{Ax}	6.54 ± 1.59 ^{Ax}	9.73 ± 3.07 ^{Ax}	2.60 ± 1.18 ^{Ax}	NS	NS	NS	NS	NS	NS
Afternoon	4.72 ± 1.85 ^{Ax}	8.77 ± 5.29 ^{Ax}	7.53 ± 2.85 ^{Ax}	2.32 ± 0.69 ^{Bx}	NS	NS	0.034	NS	0.032	0.029
Evening	5.01 ± 1.89 ^{Ax}	4.76 ± 4.32 ^{Ax}	2.84 ± 0.80 ^{Ax}	6.44 ± 2.54 ^{Ax}	NS	NS	NS	NS	NS	NS

A-C different superscripts represent statistically significant difference within the same time of day [row] and across different treatments [column] at $p<0.05$. x-y different superscripts represent statistically significant difference within the same treatment [column] and across different times of the day [rows] at $p<0.05$.

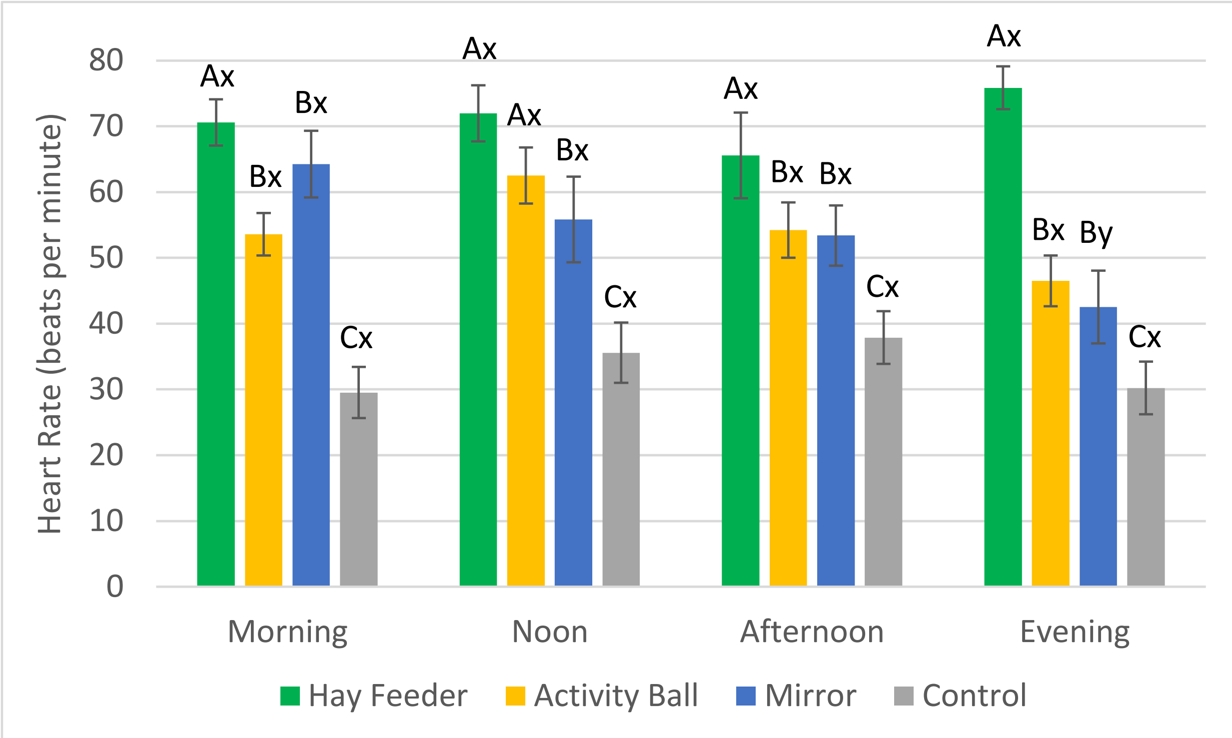


Figure 1. Heart rate (mean ± SE beats per minute) of stalled horses when provided the hay feeder, activity ball, mirror, and control during 30-minute observations at different times of the day. A-C different superscripts represent statistically significant difference within the same time of day and across different treatments at $p<0.05$. x-y different superscripts represent statistically significant difference within the same treatment and across different times of the day at $p<0.05$.

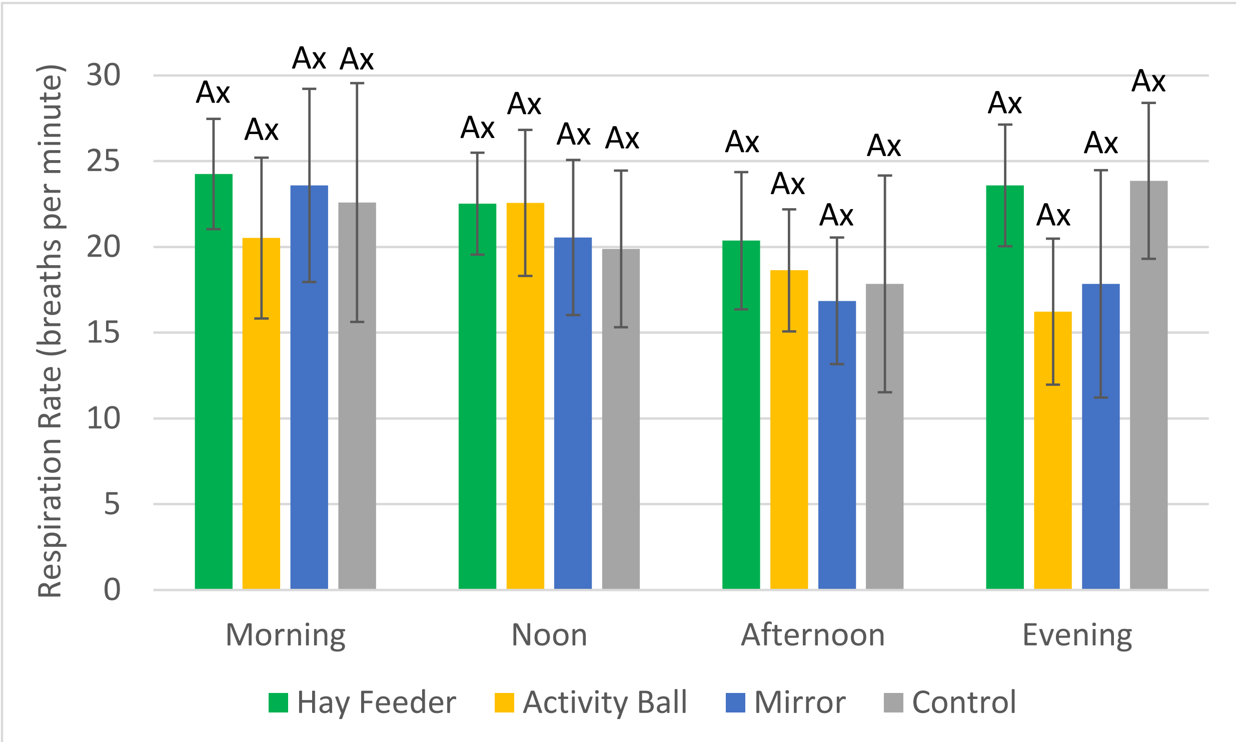


Figure 2. Respiration rate (mean ± SE breaths per minute) of stalled horses when provided the hay feeder, activity ball, mirror, and control during 30-minute observations at different times of the day. A-B different superscripts represent statistically significant difference within the same time of day and across different treatments at $p<0.05$. x-y different superscripts represent statistically significant difference within the same treatment and across different times of the day at $p<0.05$.

Figure 3. Percentage of observation (mean ± SE) the stalled horses spent performing each behavior when provided the hay feeder, activity ball, mirror, and control during 30-minute observations at different times of the day. A-C different superscripts represent statistically significant difference within the same time of day and across different treatments at $p<0.05$. x-y different superscripts represent statistically significant difference within the same treatment and across different times of the day at $p<0.05$. This figure is made up of nine charts (A through I).

Chart A: Foraging

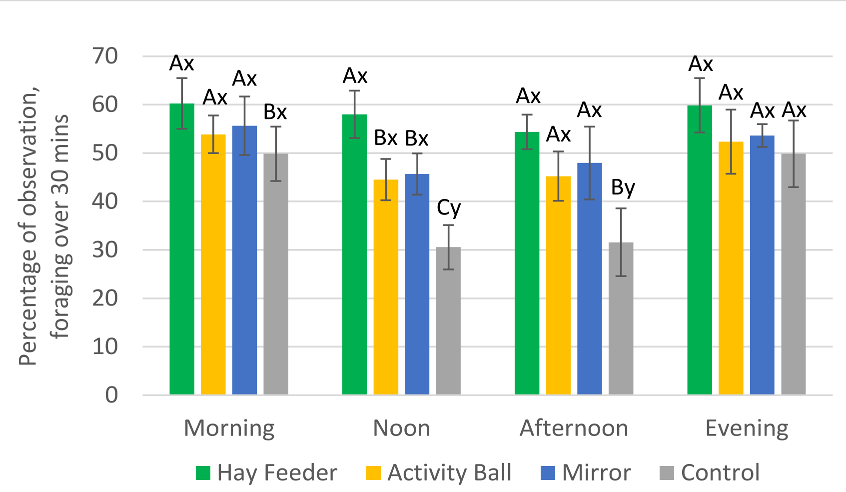


Chart B: Locomotion

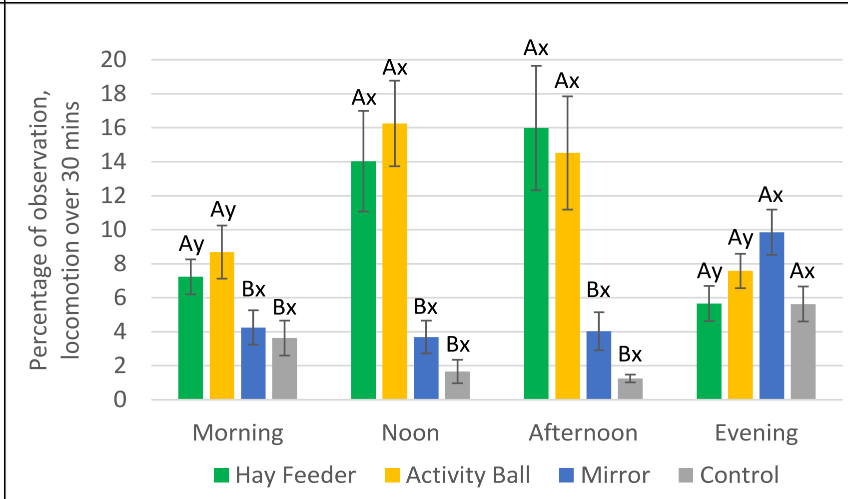


Chart C: Standing Alert

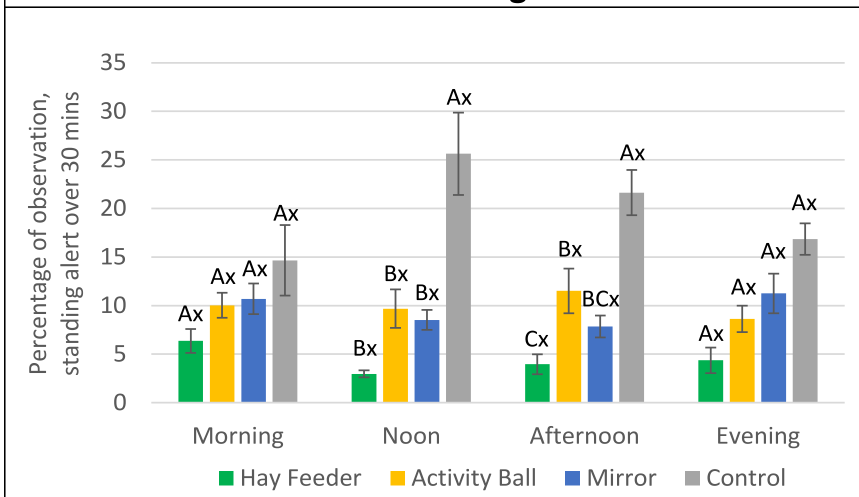
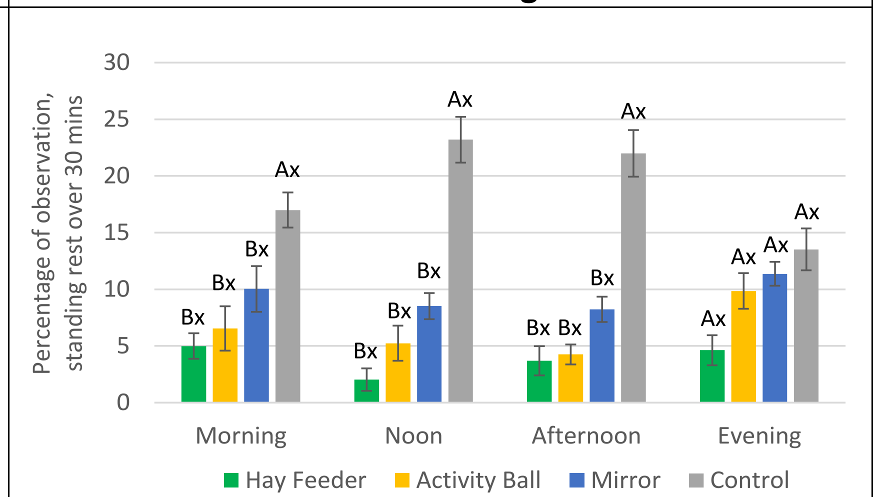


Chart D: Standing Rest



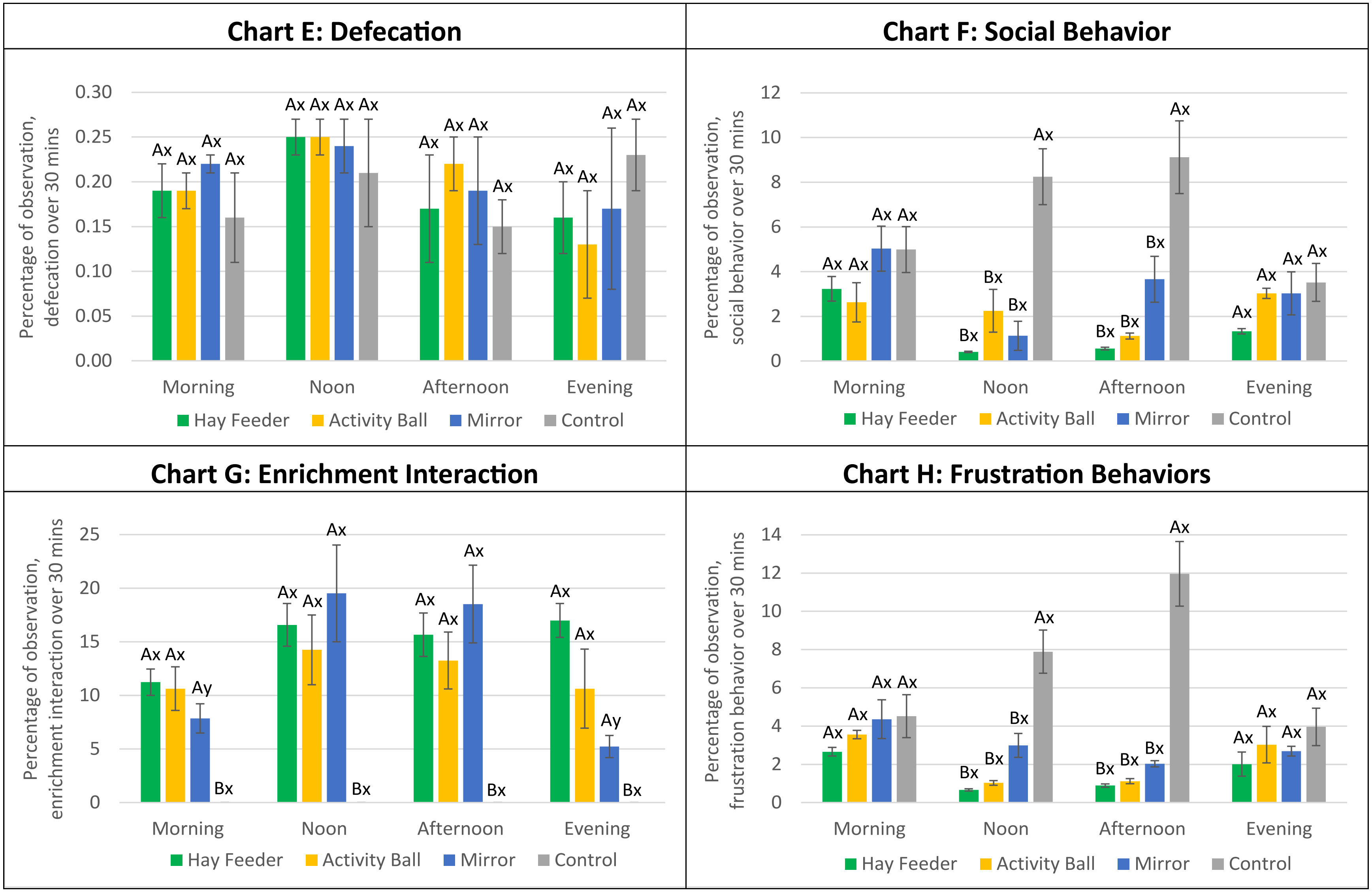


Chart E: Defecation
Percentage of observation, defecation over 30 mins
0.30
0.25
0.20
0.15
0.10
0.05
0.00
Morning
Noon
Afternoon
Evening
Hay Feeder
Activity Ball
Mirror
Control
Chart F: Social Behavior
Percentage of observation, social behavior over 30 mins
12
10
8
6
4
2
0
Morning
Noon
Afternoon
Evening
Hay Feeder
Activity Ball
Mirror
Control
Chart G: Enrichment Interaction
Percentage of observation, enrichment interaction over 30 mins
25
20
15
10
5
0
Morning
Noon
Afternoon
Evening
Hay Feeder
Activity Ball
Mirror
Control
Chart H: Frustration Behaviors
Percentage of observation, frustration behavior over 30 mins
14
12
10
8
6
4
2
0
Morning
Noon
Afternoon
Evening
Hay Feeder
Activity Ball
Mirror
Control

Chart I: Other Behaviors

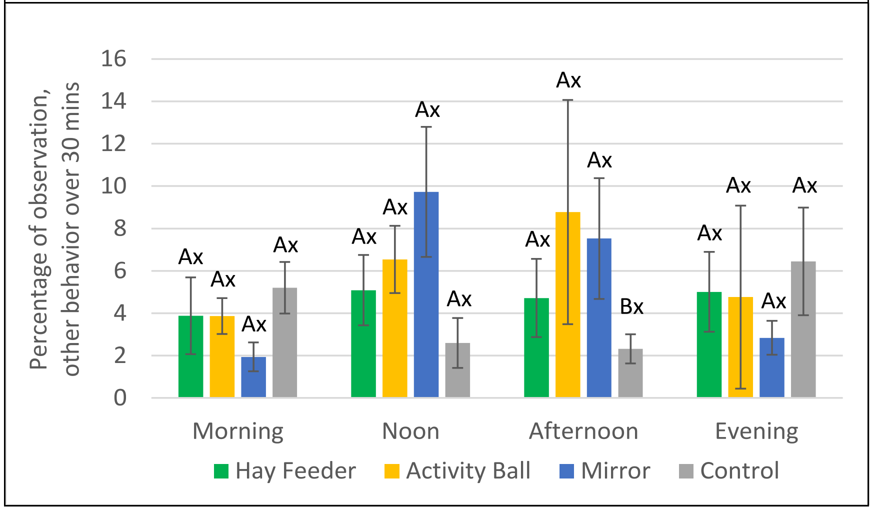


Figure 4. Activity budgets of stalled horses in the control group and when provided the hay feeder, activity ball, and mirror treatments at different times of the day. Values represent the mean percentage of observation the behavior was performed over 30-minute observations across all animals with that treatment at the associated time-of-day. For standard error variation around the means, see the associated behavior in figure 3. All three enrichment items had several significant differences compared to the control, with many of the effects being similar across enrichment items; however, there were some various effects as well. The hay feeder was the most influential among the three items at increasing foraging at noon; unlike the other items, the mirror did not have a significant effect on locomotion at any time point; and the hay feeder was more influential than the activity ball but not the mirror at reducing standing alert behavior in the afternoon. Within each treatment, there was significantly more locomotion (hay feeder and activity ball) and enrichment interaction (mirror) at noon and afternoon compared to the morning and evening, and there was more foraging (control) in the morning and evening compared to noon and afternoon. This figure is made up of twelve charts (A through P).

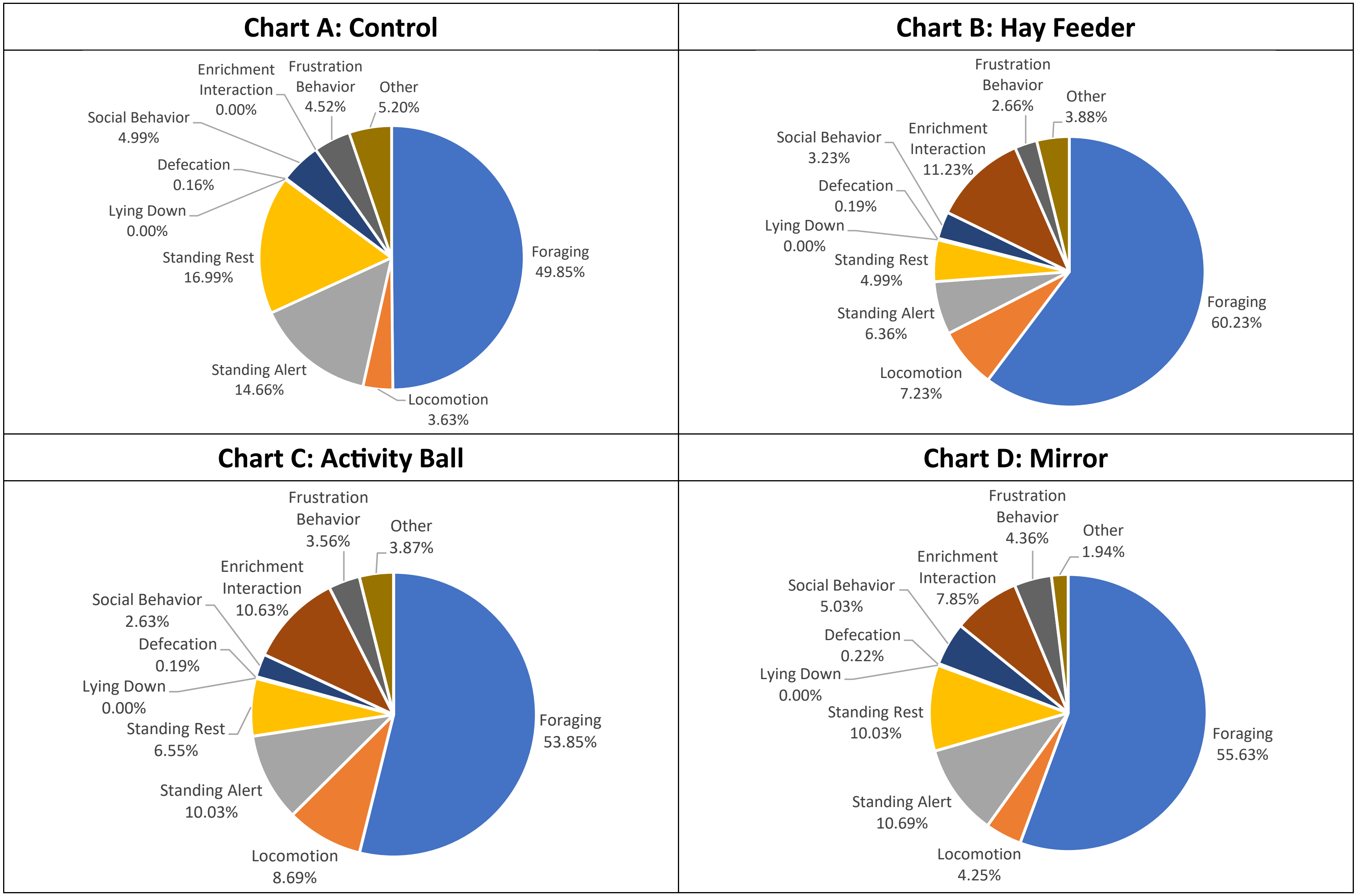


Figure 4a. Activity budgets of stalled horses in the morning when in the control group and provided the hay feeder, activity ball, and mirror.

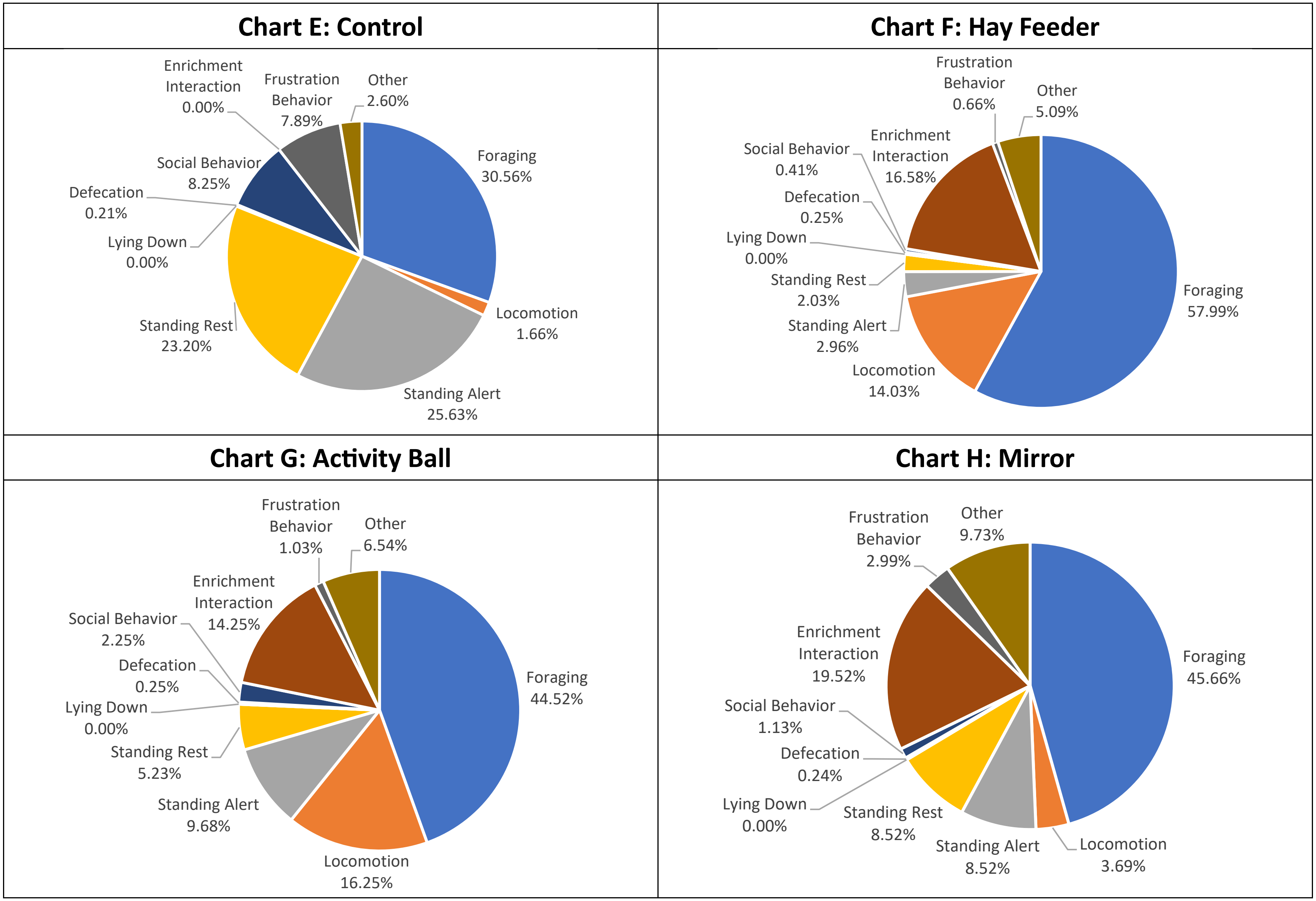


Figure 4b. Activity budgets of stalled horses at noon when in the control group and provided the hay feeder, activity ball, and mirror.

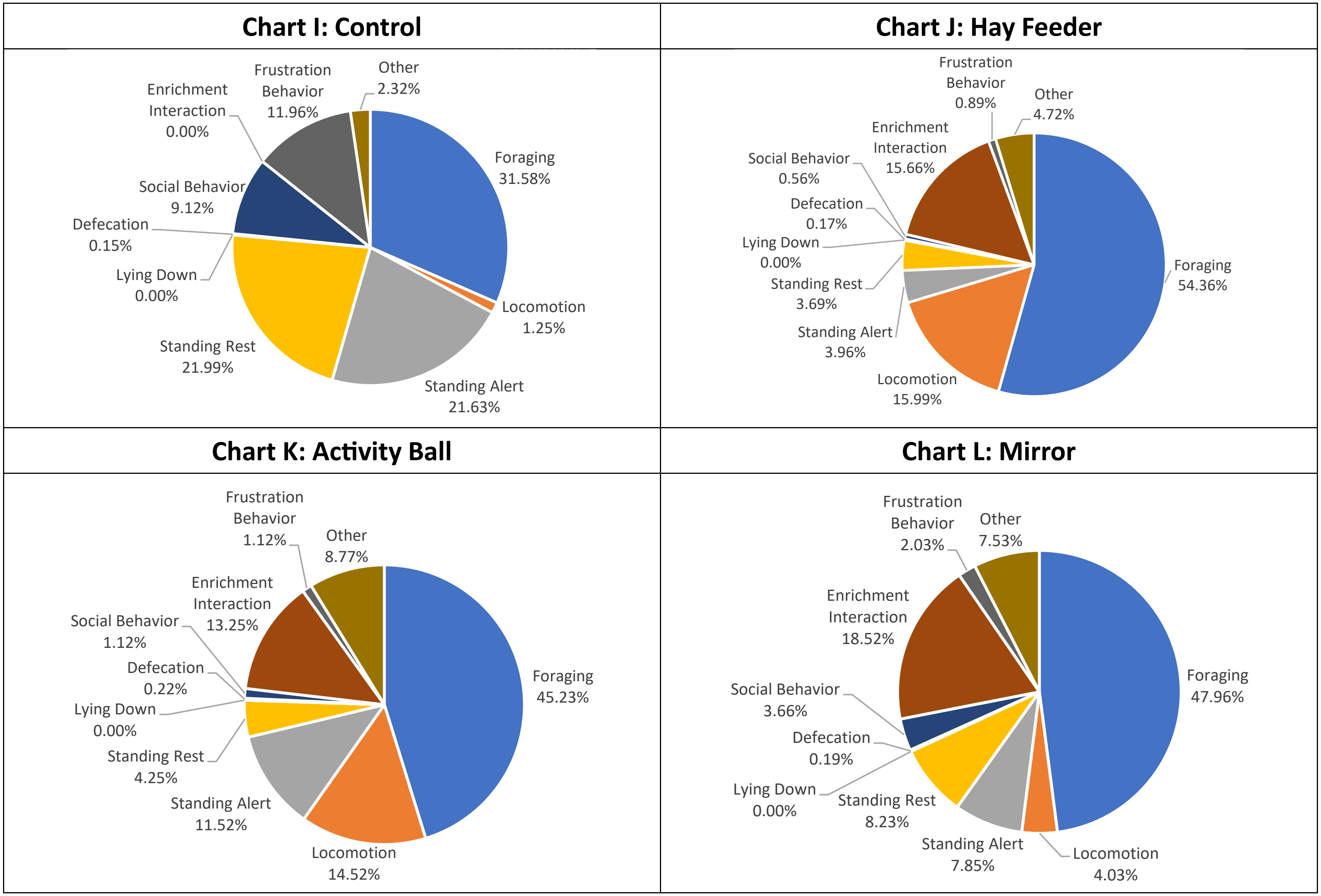


Figure 4c. Activity budgets of stalled horses in the afternoon when in the control group and provided the hay feeder, activity ball, and mirror.

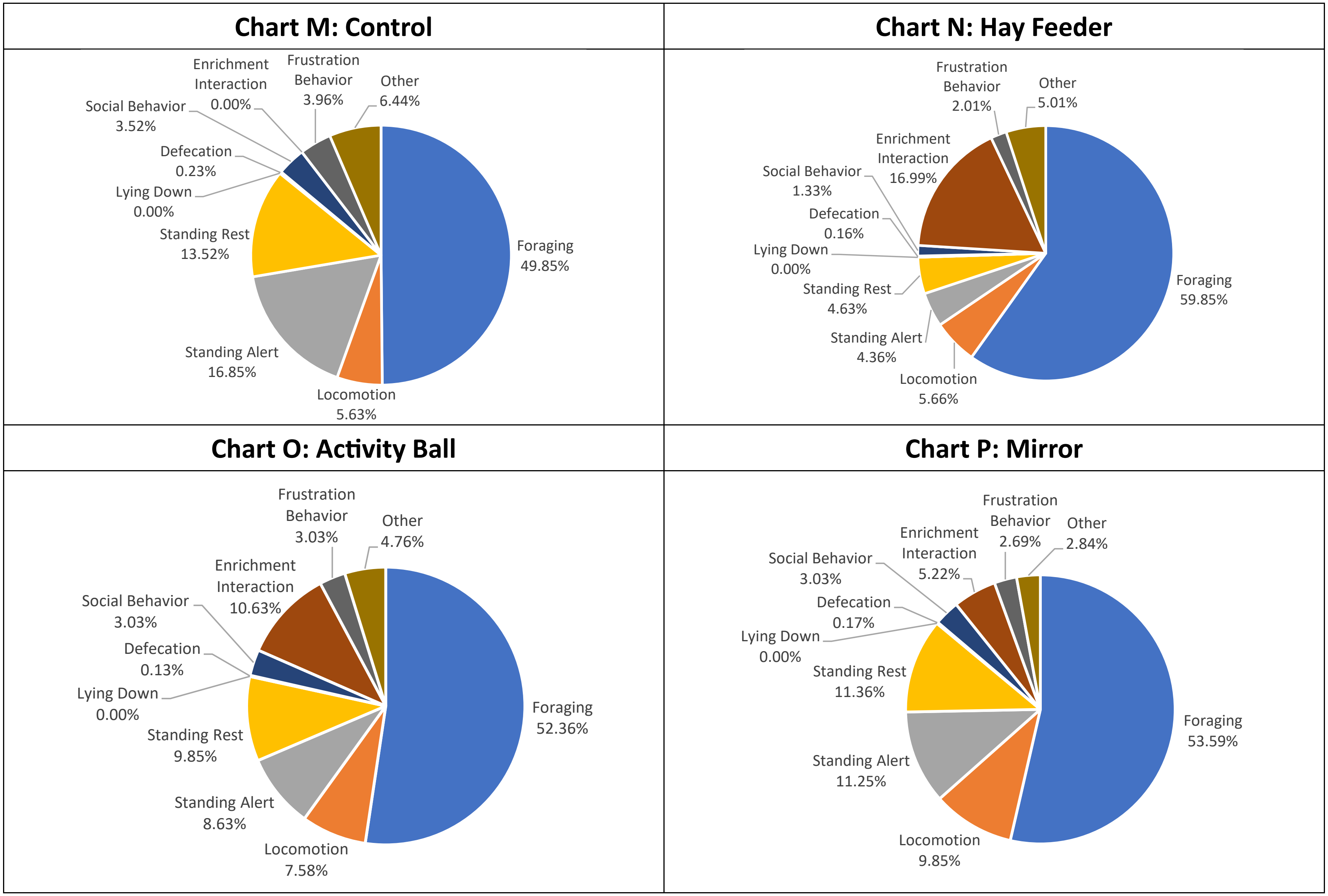


Figure 4d. Activity budgets of stalled horses in the evening when in the control group and provided the hay feeder, activity ball, and mirror.