Factors affecting the behavior of captive white rhinoceros (*Ceratotherium simum*) and the accuracy of ad-hoc keeper data

Running title: Rhino behavior and keeper data accuracy

Anne E. Goodenough[1*], Thomas W. Price[1], Danica L. Brazier[1], Katie McDonald[2]

1 = Natural and Social Sciences, University of Gloucestershire, Cheltenham, GL50 4AZ

2 = West Midlands Safari Park, Bewdley, UK

* Corresponding author

aegoodenough@glos.ac.uk +44 (0)1242 714669;

Orcid = 0000-0002-7662-6670

Abstract

Although white rhinoceros (*Ceratotherium simum*) are common in captivity, few behavioral studies have been conducted and there is seemingly no research for immersive exhibits where potential for visitor effects is high. Moreover, little information exists on possible effects of weather and temperature on rhino outside their native range. Here we analyze 14,501 observations of rhino in a drive-through enclosure. Data were collected by researchers (n=12,160 datapoints) and keepers (n=2,341 datapoints) over a four-month period. We aimed to: (1) quantify behavior using detailed researcher-collected data and contemporaneous but ad-hoc keeper-collected data; (2) compare datasets statistically; (3) establish effects of visitors, temperature, and weather on behavior; and (4) assess the influence of visitors on similarity of researcher/keeper datasets. Activity budgets were similar to the wild and the single previous study from a traditional (non-drive-through) enclosure. There was some discrepancy in activity budgets between researcher and keeper data due to significant differences in recorded frequency of two rare behaviors (horn rub; social interaction) and two behaviors that could be easily confused (grazing vs standing with head-down): recording of other behaviors matched well. Weather and temperature affected behavior, with rhino becoming more sedentary (- locomotion, grazing; + resting, standing, and sedentary eating of hay) on hot/sunny days compared to cool/wet days. The number of visitor vehicles had a fairly negligible effect but resting was lower on busy days, possibly as vigilance increased. The match between researcher/keeper datasets was lowest when visitor numbers were high, suggesting visitors might affect keeper ability to accurately record behavior.

Keywords

Activity budget, dataset comparison, drive-through enclosure, visitor effects.

Introduction

Collecting data on the behavior of animals in captivity, and how behavior changes relative to external factors, is crucial for designing suitable enclosures and informing optimal husbandry (Melfi, 2009; Whitham and Wielebnowski, 2013). Such data also allow temporal behavioral changes to be identified in response to modifications to dietary regime (e.g. Höttges et al., 2019), enclosure design (e.g. Ross, 2006; Quirke et al., 2012) and enrichment (e.g. Coelho et al., 2012; Vaz et al., 2017). Monitoring behavior can also allow detection of change in the health or welfare status of individual animals (Rose and Riley, 2019). A considerable amount of research in recent years has been dedicated to the effects of visitors on animal behavior (Hosey, 2000; Davey, 2007; Fernandez et al., 2009) and how visitor effects co-vary with factors such as weather and time of day (Goodenough et al., 2019; Sherwen and Hemsworth, 2019). Visitor effects are both species- and situation-dependent. In some cases, visitor presence has no substantive effect on behavior (e.g. Margulis et al., 2003; Goodenough et al., 2019) or can be beneficial by acting as a stimulus (e.g. Choo et al., 2001; Jones et al., 2016). However, sometimes visitor presence can decrease social or maintenance behaviors in captive animals (Wood 1998; Chamove et al., 1988) or increase non-desirable or vigilance behaviors (Birke 2002; Blaney and Wells 2004; Kuhar 2008; Larsen et al., 2014; Sherwen et al., 2015).

The southern white rhinoceros (*Ceratotherium simum simum*) is the most common (sub)species of rhinoceros in captivity: there are ~650 individuals in zoological collections worldwide. Husbandry is supported by comprehensive Best Practice Guidelines by AZA in America (Fouraker and Wagener, 1996) and EAZA (Versteege, 2018) in Europe. Although rhino appear to adapt well to captivity (Hutchins and Kreger, 2006), comparatively little detailed behavioral research has been undertaken. One initial study documenting the activity budget of 14 southern white rhino housed in a traditional enclosure at a UK zoo showed very similar patterns to free-ranging animals (O'Connor, 1986, and Owen-Smith, 1973, respectively). The main difference was the influence of weather: in wild animals temperature correlated negatively with activity while cloud cover correlated positively with activity (Owen-Smith, 1973) but these relationships were not seen for captive animals (O'Connor, 1986). Somewhat surprisingly, there has been no published research on the behavior of white rhino housed in drive-through enclosures, even though this type of enclosure is commonly-used for white rhino (Versteege, 2018). This means that the possible effects of visitors on behavior in immersive exhibits have not been considered, despite potential for visitor effects being high in such settings (Goodenough et al., 2019). Research that further illuminates rhino behavior, and external influences on behavior, could be useful especially given the intense pressures on the species in the wild (e.g. Penny et al., 2019).

One of the barriers to undertaking behavioral research on captive animals – and a key reason for substantial gaps in zoo husbandry knowledge such as those identified here for white rhino – is the amount of data required to make evidence-informed decisions (Melfi, 2009). Although collecting behavioral data via ethograms is conceptually straightforward (Lehner, 1998), behavior of zoo animals can be influenced by temporal variables such as time of day (Grandia et al., 2001; Maia et al., 2012), abiotic factors such as temperature and precipitation (Rees, 2004; Bouchard and Anderson, 2011; Young et al., 2012), visitor numbers (Hosey, 2000; Davey, 2007; Fernandez et al., 2009) and variation in age, sex, origin, rearing and personality (Melfi, 2009; Whitham and Wielebnowski, 2013). Obtaining suitable and sufficient data therefore involves considerable time and effort.

Theoretically, the people best-placed to collect behavioral data are keepers. Keepers can usually identify individual animals using variation in size or markings, and typically spend a considerable amount of time within or adjacent to the enclosure (Less et al., 2012; Carlstead et al., 2019). Some published research has been used keeper data where the focus has been animal personality using subjective scoring of traits such as boldness (e.g. Grand et al., 2012; Yasui et al., 2013). However, most systematic behavioral data are collated by dedicated researchers (staff or research students) rather than keepers. The fact that the people seemingly best-placed to collect the data rarely actually do so is driven several issues. Firstly, keepers are not always trained in formal data collection methods and may not be viewed as being part of a research team. Even where formal behavioral data have been collected by keepers for published research, ethograms have usually been simplified. This reduces inter-observer variation but means that detailed understanding of behavior is often lacking (Gosling, 2001; Less et al., 2012). Secondly, keepers (or their managers) might not be motivated to participate in research, possibly because they do not see tangible on-ground changes in practice. Thirdly, and perhaps most importantly, keepers do not usually have much time to devote to data collection (Kuhar, 2006; Less et al., 2012). In addition to a general lack of time to undertake activities not primarily related to husbandry, keeper availability is often constrained temporally (e.g. only possible at the start or end of the day), infrequent (e.g. a scan sample when a keeper walks past the exhibit) or ad-hoc (e.g. collected whenever the exhibit is quiet). Even when data collection is possible this might involve keepers multitasking, for example, while ensuring visitor safety. Keeper-collected datasets are thus often less detailed than researcher-collected data, with fewer datapoints per day and smaller overall sample sizes. This means that rare behaviors can be under-recorded (Kuhar, 2006). More importantly, if there are systematic differences in keeper availability throughout the day and also systematic differences in behaviors throughout the day, keeper data might not be fully representative of the overall activity budget (Carlstead et al., 2009). There is also the possibility that external factors such as visitor numbers – itself often a key determinant of animal behavior – might co-vary with keeper ability to collect accurate data due to competing time demands when visitor numbers peak. It is thus vital to cross-validate ad-hoc keeper-collected data with detailed researcher data before the former are relied upon (Gosling, 2001; Carlstead et al., 2009; Less et al., 2012). A better understanding of the circumstances in which keeper-collected data might be more or less reliable would also be valuable.

In this paper, we quantify the behavior of rhino housed in a drive-through enclosure using detailed researcher-collected data and contemporaneous ad-hoc keeper-collected data. Our first aim is to generate a group-level rhino activity budget using ad-hoc keeper data and detailed researcher data separately to allow overall similarity to be assessed graphically. Our second aim is to statistically quantify the effects of dataset, the interaction between dataset and recording day, and the interaction between dataset and rhino ID, on rhino activity budget. These models will not only allow us to establish whether any differences between datasets are significant, but also to determine whether multi-day datasets are needed to mitigate temporal variations in data accuracy and whether data are equally robust for all animals. We will then extend this by comparing the frequency occurrence of individual behaviors between datasets. Once datasets have been compared, our third aim is to investigate potential impacts of visitor numbers, temperature, and weather on rhino behavior. Finally, our fourth aim is to consider whether visitor numbers affect the similarity between researcher and keeper data, which might occur if ability of keepers to collect accurate behavioral data is reduced due to increased safety patrols when visitor numbers peak.

Methods

Study setup

This research was conducted at West Midland Safari Park (WMSP, Worcestershire, UK) in Autumn 2019, where we studied eight southern white rhinoceros: one adult male, four adult females, one sub-adult male (4 years old), one juvenile female (23 months old), and one juvenile male (16 months old). Rhino were released from their overnight accommodation into an "African Plains" exhibit about 10:00-10:30 hrs and were returned to their overnight accommodation about 15:30-16:00 hrs. The African Plains exhibit covered around 8 ha, dominated by grassland bisected by vehicular roads long which visitors drove their own vehicle. Hay feeders, browse feeders, mud wallows, wooden shelters and water troughs were distributed throughout the enclosure. The exhibit was enclosed by wood and chain-link fencing, except along a shared boundary with an African lion (*Panthera leo*) exhibit where solid wooden panels were used. Cattle grids prevented animals from leaving the enclosure via roads where vehicles entered and exited. Vehicle flow was regulated by gates between the African Plains enclosure and the rest of the safari park. Rhino shared the African Plains enclosure with northern giraffe (*Giraffa camelopardalis*), plains zebra (*Equus quagga*), common eland (*Taurotragus oryx*), Congo buffalo (*Syncerus caffer nanus*), common waterbuck (*Kobus ellipsiprymnus*) and red lechwe (*Kobus leche*). During opening hours, keepers were present in patrol vehicles at all times to monitor both animals and visitors, replenish hay and browse feeders, and carry out general maintenance.

Behavioral data

We collected data when rhino were in their daytime enclosure (5-6 hrs per day) using an ethogram and instantaneous scan sampling (Lehner, 1998). Data were collected at individual level, with each rhino being identified using differences in horn morphology, body size, genitalia and density of hair on ears and tail. The initial ethogram listed 23 mutually-exclusive behaviors within 8 overarching behavioral categories, plus "out-of-sight" (Table 1). The ethogram was informed by pilot studies and consultation of literature, notably Owen-Smith (1973) and O'Connor (1986), and EAZA Best Practice Guidelines (Versteege, 2018). Before the study, ethogram definitions were agreed upon by all those collecting data (researchers and keepers). Then, to reduce inter-observer variability, a practice session was held to check that observers watching the same rhino at the same time were interpreting rhino behavior in the same way and using the same ethogram category when recording this. Collection of data was entirely non-interventional, involving researchers and keepers passively observing rhino from a vehicle that was parked at a distance from the animals. This survey protocol was considered by the screening policy of both University of Gloucestershire and West Midland Safari Park, which concluded that formal ethical approval was not necessary.

Two datasets were collected over a four-month period between September and December 2019:

- Researcher data: behaviors were recorded for each of the 8 rhino individually using instantaneous scan samples once every 3-4 minutes for 23 non-consecutive days. This gave approximately 80 observations (scans) per rhino per day, totaling 12,160 observations across the entire data collection period with an average of 1,520 observations per rhino. These data were collected by co-authors TP and DB, who were solely focussed on research data collection.
- Keeper data: data were collected on the same days as the researcher data. Our aim was for keepers to collect data on each rhino using an instantaneous scan sample once every 20 minutes (i.e. 15-18 records per day depending on exact timing of rhino release from, and return to, their overnight accommodation). The intention was thus that keeper data would be matched to the researcher data in terms of data collection method, collection days, and start/finish times, with the only difference being the frequency of sampling and thus the overall daily sample size. However, as expected given the other roles that keepers were performing, data collection was not always possible. This often meant that either the time between scan samples often increased beyond the expected 20 minutes or that sample points were completely missed, especially towards the end of the day. The resultant data were thus rather irregular and ad-hoc, with 2,341 observations in total (range = 7-18 observations per rhino per day; mean = 12.7 observations per rhino per day). Keeper data were collected by seven keepers, all of whom were familiar with the rhino, with recording performed by whomever was on duty at the relevant sample point.

After data collection, the researcher dataset was condensed by summing observations to give one line of data per rhino per day; the same process was repeated separately for the keeper data. This allowed the datasets, which differed in sample size, to be directly compared and avoided pseudo-replication (i.e. multiple lines of data per rhino per day, which would have been non-independent). In total, there were 179 "rhino days" within each dataset (23 days * 8 rhino, minus 5 days on which different specific rhino were not released into the outdoor enclosure). Non-release of specific rhino occurred for husbandry-related reasons. The number of behaviors was also reduced, with behaviors being retained, combined, or removed from the dataset depending on the frequency of observations and behavior distinctness. This reduced the number of behaviors from 23 to 14. This process is detailed fully in Table 1, briefly: two behaviors that were never witnessed (chasing and mouthing of genitals) were removed, two types of horn aggression were combined, urination and defecation were combined, all three witnessed reproductive behaviors were combined, and the rarely-observed behaviors of sniffing, vocalizing, drinking and body rubbing were combined into "other".

The number of out-of-sight records per rhino per day was recorded on both datasets to allow this to be directly compared. However, when determining activity budgets and undertaking detailed behavioral analysis, out-of-sight records were removed to ensure that the percentage of each behavior was not skewed by differences in out-of-sight observation frequency when comparing rhinos, days, and datasets.

Non-behavioral data

In addition to behavioral data, we recorded: (1) daily total number of visitor vehicles entering the safari park; (2) weather ranked an ordinal scale (1 = heavy rain; 2 = light or intermittent rain; 3 = overcast; 4 = sunny intervals with some cloud; 5 = sunny); and (3) average temperature (°C). Visitor vehicle data were used rather than number of visitors *per se* since the number of vehicles was more relevant to a drive-through enclosure (daily min = 152 vehicles, daily max = 1,131 vehicles, daily mean = 315 vehicles). Temperature data were collected hourly across the 5-6 hours of data collection and averaged to give a single figure per day (min = 3.0°C, max = 22.3°C, mean = 12.9°C).

Data analysis

Activity budgets were visualized using stacked histograms. One histogram was created using researcher data from all rhinos across the entire study period; this process was also undertaken separately using the keeper data. To summarize behavior in one overarching composite variable, and thus give a single proxy measure for activity budget per rhino per day per dataset, Principal Components Analysis was used. This used percentage data rather than original frequencies to standardize data relative to the very different sample sizes (~80 observations per rhino per day for researcher data *versus* ~13 observations per rhino per day for keeper data). The first Principal Component (PC1), which was normally distributed, explained 61.4% of the variation in the underpinning data, with the main loadings being grazing (-0.946), as well as standing (+0.652), sedentary eating of hay (+0.545) and walking (+0.430). PC1 was then used as the dependent variable in subsequent analysis described below. Only PC1 was modelled since the intention was to analyze patterns in the overall activity budget (where PC1 was the most appropriate single proxy variable) prior to subsequent analysis on specific behaviors using raw data.

The first General Linear Model was used to quantify whether activity budget was significantly related to: (1) day, (2) dataset, and (3) the interaction between day and dataset. The second General Linear Model was used to quantify whether overall activity budget was significantly related to: (1) rhino, (2) dataset, and (3) the interaction between rhino and dataset. Between them, these two models allowed consideration of whether overall activity budget varied significantly on different days, whether it was significantly different between rhino, whether there was a significant difference between the datasets overall, and whether the dataset similarity varied on different days or for different rhino. While it would not necessarily be surprising if activity budget differed between rhino or between days *per se*, the rationale for this modeling was to calculate the interaction term to assess whether datasets similarity differed temporally (and thus whether multi-day datasets would be needed to mitigate temporal variations in data accuracy) and whether the datasets where more similar for some rhino than others (and thus whether keeper data accuracy would be higher for some animals than others). Two 2-way models were required as there were insufficient replicates for a single 3-way model (there was one line of data per dataset per rhino per day so dataset, rhino ID, and day could not be analyzed in a single model). Then, to compare the relative frequency occurrence of individual behaviors between the two overall datasets statistically, Z tests for proportions were performed using the relevant numerator and denominator values (i.e. total frequency of each behavior in the different datasets relative to the total number of datapoints in those datasets after out-of-sight records had been excluded). This test was appropriate because it does not assume a normal distribution, instead being based on proportions calculated from binomial data. The comparison of proportions rather than raw data meant that difference in sample sizes between datasets was not problematic.

To establish effects of daily visitor vehicle numbers, mean daily temperature, and daily weather on rhino behavior, General/Generalized Linear Models were undertaken on the researcher dataset. The first model used PC1 as a proxy for overall activity budget as the dependent variable; subsequent models were created for grazing, standing, walking, sedentary eating of hay, resting, running, and social interaction (the seven behaviors most commonly observed; see results). When modeling the normally distributed and continuous PC1, a General Linear Model with a normal error distribution with an identity link function was used. When modeling specific behaviors using percentage data, Generalized Linear Models were used with a binomial error distribution and a logit link function.

Finally, Generalized Linear Models were conducted to establish the effect of visitor vehicle numbers on the similarity between researcher and keeper datasets as regards recording of specific behaviours. The same suite of seven behaviors was tested as used as above (grazing, standing, walking, sedentary eating of hay, resting, running, and social interaction). In all cases, the absolute difference between the percentage scores for the relevant behavior per rhino per day was calculated for the two datasets (i.e. if one dataset had a score of 50% for grazing for a specific rhino on a specific day and the equivalent score of the other dataset was 60%, the absolute difference was 10 regardless which dataset was the higher). In this way, small numbers indicated greater similarity between the datasets. Difference scores were not normally distributed so models were fitted with a Poisson error distribution and a log link function.

Results

Rhino activity budgets

Rhino activity budgets were dominated by four commonly-occurring behaviors in both researcher and keeper datasets: grazing, standing, walking, and sedentary eating of hay (Fig. 1). In the researcher dataset, these behaviors accounted for 94.8% of observations compared to 92.8% of observations in the keeper dataset. The remaining nine behaviors occurred less frequently, with most of these individually accounting for <1% of the activity budget (Fig. 1). In total, rhino were classified as out-of-sight in 15.7% of researcher observations, compared to just 0.2% for keeper observations. However, while researchers collected data for 100% of expected sample periods, keepers missed 20.8% of sample periods because they were unable to record data at that time. The missing keeper data was sometimes due to keepers being unavailable and sometimes because rhino were not observable (i.e. out-of-sight); these could not be decoupled from the data submitted. Anecdotal evidence also suggested that keepers might be unwilling to record "out of sight" in case this was regarded as poor working practice.

Comparison of detailed research data with contemporaneous ad-hoc keeper data

The first General Linear Model used PC1 to summarize the overall activity budget within a single variable and showed significant effects for day and dataset and the interaction between these two factors (overall model: $f = 9.200$, d.f. = 45, $p < 0.001$; day: $f = 15.968$, d.f. = 22, $p < 0.001$; dataset: $f = 3.996$, d.f. = 1, $p = 0.046$; day*dataset interaction: $f = 2.686$, d.f. = 22, $p < 0.001$). The second General Linear Model on PC1 showed significant differences for Rhino ID and dataset but no significant interaction (overall model: $f = 2.804$, d.f. = 15, $p < 0.001$; rhino ID: $f = 5.326$, d.f. = 7, $p < 0.001$; dataset: $f = 3.996$, d.f. = 1, $p = 0.046$; rhino*dataset interaction: $f = 0.409$, d.f. = 7, $p = 0.897$). Taken together, these two models showed – perhaps unsurprisingly – that activity budgets varied between days and between rhino. More importantly in the context of this study, however, the models also confirmed: (1) that quantification of rhino activity budget differed between researcher and keeper data (dataset significant in both models); (2) that on some days the datasets were more similar than others (significant day*dataset interaction); but (3) that the similarity between datasets did not differ for specific rhino (non-significant rhino*dataset interaction). These interactions are important because they indicate that multi-day keeper data are needed as results from a single day might not be representative but that keeper data are equally robust for all animals rather than being biased towards specific animals.

We then explored the frequency occurrence of individual behaviors between the datasets using Z test for proportions. The majority of behaviors did not differ significantly between datasets ($p \geq 0.05$). The exceptions were: grazing ($z = 6.987$; $p < 0.001$, higher in researcher data), standing ($z = -7.567$; $p < 0.001$, higher in keeper data), and social interaction ($z = -5.246$; $p < 0.001$, higher in keeper data). The differences for grazing and standing were particularly important given these were common behaviors.

Horn rubbing was the only behavior recorded in one dataset (researcher) not recorded in the other (keeper). When individual behaviors recorded for all rhino across the entire study period were grouped into the overarching behavioral categories (Table 1), agreement between datasets was high, especially for common behaviors (Fig. 2).

Effect of visitors, temperature and weather on rhino behavior

General Linear Modeling of the researcher dataset showed that overall activity budget (as summarized by PC1) was significantly affected by temperature and weather; the number of visitor vehicles were non-significant (Table 2). Follow-up Generalized Linear Modeling on frequency occurrence of specific behaviors (also reported in Table 2) revealed that grazing, standing and walking behaviors all occurred more often when the weather was wet/overcast, while sedentary eating of hay occurred more often when the weather was dry/sunny. Grazing and walking were exhibited more frequently at lower temperatures, while eating hay and resting were exhibited more frequently at higher temperatures. Resting was less frequent, but standing was more frequent, when visitor numbers were high.

Effect of visitors and weather on similarity between researcher and keeper datasets

The similarity between researcher and keeper data was only significantly related to the number of visitor vehicles for two behaviors: resting ($\chi^2 = 17.103$; $p < 0.001$) and standing ($\chi^2 = 19.081$; $p < 0.001$): for both behaviors the datasets became less well-matched when visitor numbers were high. This is important given that resting and standing were also the only two behaviors that were significantly influenced by visitor vehicle numbers (Table 2). Standing was also one of three behaviors where there was an overall difference between researcher and keeper datasets. Resting did not differ overall between researcher and keeper datasets, despite the divergence of the scores when visitor numbers were high, possibly because the overall frequency of this behavior was comparatively low overall.

Discussion

Our study has indicated that the three most prevalent diurnal behaviors in a group of captive rhino, which were housed in a drive-through enclosure and observed between September and December, were grazing, standing, and walking. This agrees with a previous study of rhino in a traditional enclosure by O'Connor (1986). However, we demonstrated that overall rhino activity budgets, and the frequency of specific behaviors, is influenced both by temperature and weather. These patterns make intuitive sense: rhino were more sedentary on hot/sunny days (increases in resting, standing, and sedentary eating of supplemental hay) while walking and grazing (which occurs as animals move) decreased. With the exception of eating supplemental hay, these patterns mirror those found for wild rhino where higher temperatures and low cloud cover decreased grazing and increased resting (Owen-Smith, 1973). Such effects have not been previously recorded for captive rhino (O'Connor, 1986).

Visitor numbers, as measured by the number of vehicles passing through the drive-through enclosure, did not affect the overall activity budget. However, large numbers of visitors were associated with changes in two specific behaviors: decreasing resting and increasing standing. As both these behaviors are inactive, visitor numbers did not alter the relative proportion of time spent active/inactive, but it is possibly reflective of a general increase in vigilance or "readiness to flee". This type of response to high visitor numbers has been seen previously in orangutans (*Pongo pygmaeus*) (Birke, 2002), koala (*Phascolarctos cinereus*) (Larsen et al., 2014) and little penguins (*Eudyptula minor*) (Sherwen et al., 2015). Overall, however, we conclude that the effects of visitors on rhino in this setting seem to be minimal. However, it is recommended that rhino are monitored if there are multiple consecutive high-visitor days to ensure that increased vigilance behavior are not associated with physiological effects such as elevated stress hormones. Such responses have been observed previously for rhino in stressful situations in the wild (Penny et al., 2020) and for other pachyderms in captivity (Grand et al., 2012).

Agreement between detailed researcher and ad-hoc keeper datasets was high for most behaviors, suggests that using keepers to collect baseline data to monitoring management and husbandry is appropriate. This agrees with previous work by Less et al. (2012), where keeper assessments of western lowland gorilla (*Gorilla gorilla gorilla*) activity levels were validated against detailed behavioral data collected systematically. However, there were some important significant differences between datasets, including for two common behaviors. Grazing was recorded statistically more often, and standing was recorded statistically less often, in researcher data relative to keeper data. A possible explanation is that, because the typical rhino standing posture is head-down given head weight (unless the animal is alert/vigilant), head-down grazing and head-down standing were confused. Confusion of behaviors that share superficial similarities has been seen previously in ad-hoc data: Williams et al. (2012) found that citizen scientists observing Asian short-clawed otters (*Aonyx cinereus*) confused playing [in the water] and swimming and the frequency of these two behaviors differed to contemporaneous researcher data until they were combined. In our rhino data, the differences between researcher and keeper data for grazing and standing cancelled one another out when these behaviors were combined. Because grazing and standing were commonly observed and thus a key element of the overall activity budget, differences in the frequency recording of these two behaviors was the driver for the statistically significance difference in overall activity budget (as summarized by PC1).

As expected, there were also differences between datasets for uncommon behaviors between detailed researcher and ad-hoc keeper data. Horn rubbing (which is known to be rare even for wild rhino and only apparent in detailed datasets that either involve numerous scan samples over a prolonged period or continuous sampling (Penny et al. (2021)), was recorded as a rare behavior by researchers but absent from keeper data. This highlights the importance of having detailed data with short times between scans, as found previously by Orban et al., (2016) when recording rare behaviors such as pacing (1.4% prevalence) in Giraffe. Frequency of social interaction, another uncommon behavior, was also recorded at significantly different rates but the direction of the difference was unexpected as keepers recorded social interaction proportionally more often than researchers despite the sampler sample size and greater time interval between scans. It is possible that keepers, who know each rhino and their specific personalities well, might have detected subtle forms of social interaction (e.g. body language), which were missed by the researchers. Interestingly, it was also notable that there was greater variation in the keeper data for socializing (a behavioral category that contained the single behavior of social interaction) evident in Fig. 2. This difference in variability could be due to the inherent smaller sample sizes in keeper data, the effect of inter-observer variation in keeper data (multiple keepers were involved in data collection), or a combination. Potential variation in interpreting subtle behaviors between people would be an avenue worthy of future research so that additional calibration and training can be given if necessary (Less et al., 2012). More generally, it should be noted that instantaneous scan sampling (whether systematic and regular or ad-hoc and irregular) is not always ideal for recording behaviors that are rarely exhibited and/or short in duration when a continuous recording method, such as all-occurrence sampling. is preferable (Lehner, 1996). This has been found when surveying Gorilla where rare short-duration behaviors such as hand clap and chest beat could only be recorded reliably using continuous sample (King et al., 2003) and for surveying of Black-faced Spoonbills (*Platalea minor*) where accurate recording of rare alarm and social behaviors was only possible using continuous sampling (Choi et al., 2007).

Our conclusions and recommendations are:

- Ad-hoc keeper data can constitute a valuable source of information for baseline activity budgets and routine monitoring of animal behavior and the ability of keepers to collect data that make a real difference to informing husbandry should be recognized. However, the importance of recording when animals are out of sight – rather than simply not collecting data – should also be emphasized, with keepers being reassured this is inevitable in ethological research rather than being any reflection on their skills. This is vital to ensure that keepers are not inadvertently biasing activity budgets, especially in cases where animals that are out of site are likely to be engaging in specific behaviors (e.g. using indoor enclosures to rest).
- It is important to consider whether there is systematic bias in the reliability of keeper data for different animals, on different days, or in relation to key external variables such as visitor numbers as occurred here. In particular, if there is an underlying bias in data accuracy in relation to visitors, this has consequences for research on visitor effects as inaccuracy in keeper data could lead to important impacts of visitors on behavior being missed (or, alternatively, "patterns" arising that are artefacts of the circumstances under which data were collected). Where necessary, care should be taken to ensure if the data recorder can be seen by the animals being studied, keeper presence does not influence the behavior of animals.
- Where detailed understanding of behavior is needed, and especially if rare or short-duration behaviors are important, consideration should be given to using detailed researcher data and ideally collected contentious recording methods (Lehner, 1996; King et al., 2003; Choi et al., 2007). If instantaneous sampling is used, large sample sizes and short durations between scans are recommended (Orban et al., 2016).

- Even where activity budgets using keeper data are largely reflective of activity budgets using researcher data when compared visually, statistical analysis is needed to empirically test whether the recorded frequency of specific behaviors differs (Gosling, 2001; Carlstead et al., 2009; Less et al., 2012). This is especially important when individual behaviors are of special interest (e.g. non-desirable or stereotypic behaviors; reproductive behaviors). It is also important when activity budgets are dominated by a few key behaviors that might be easily confused since mismatch in these can skew an entire activity budget, even when there is close agreement in the recording of most behaviors, as occurred here.
- It is vital when possible visitor effects are being assessed that other external factors that co-vary with visitor numbers and that themselves might affect animal behavior themselves, such as weather and temperature, are also analyzed using a multivariate framework. This is important to stop visitor effects being over-estimated (Goodenough et al., 2019).

Acknowledgements

We gratefully acknowledge the support of West Midland Safari Park Head of Wildlife, Angela Potter, and Head Keeper of Hoof Stock, Lisa Watkins, for their support of this project. We thank all the rhino patrol keepers from West Midland Safari Park for their help with data collection over the study period. Special acknowledgement is due to keeper, Becky Nock, whose Diploma in Management of Zoo and Aquarium Animals project initially inspired this research.

Conflict of Interests

The authors declare no conflict of interests.

Data availability statement

The data that support the findings of this study are available from the corresponding author upon reasonable request.

References

Birke, L. (2002). Effects of browse, human visitors and noise on the behavior of captive orang-utans. *Animal Welfare*, 1, 189-202.

Blaney, E. C., & Well,s D. L. (2004). The influence of a camouflage net barrier on the behavior, welfare and public perceptions of zoo-housed gorillas. *Animal Welfare*, 13, 111–118.

Bouchard, L. C., & Anderson, M. J. (2011). Caribbean Flamingo resting behavior and the influence of weather variables. *Journal of Ornithology*, 152, 307-312.

Carlstead, K., Mellen, J., & Kleiman, D. G. (1999). Black rhinoceros (*Diceros bicornis*) in US zoos: I. Individual behavior profiles and their relationship to breeding success. *Zoo Biology*, 18, 17-34.

Carlstead, K., Paris, S., & Brown, J. L. (2019). Good keeper-elephant relationships in North American zoos are mutually beneficial to welfare. *Applied Animal Behavior Science*, 211,103-111.

Chamove, A. S., Hosey, G. R., & Schaetzel, P. (1988). Visitors excite primates in zoos. *Zoo Biology*, 7, 359–369.

Choi, C. Y., Nam, H. Y., & Lee, W. S. (2007). Measuring the behaviors of wintering Black-faced Spoonbills (*Platalea minor*): comparison of behavioral sampling techniques. *Waterbirds*, 30, 310-316

Choo, Y., Todd, P. A., & Li, D. (2011). Visitor effects on zoo orangutans in two novel, naturalistic enclosures. *Applied Animal Behavior Science*, *133*, 78-86.

Coelho, C. M., Schetini de Azevedo, C., & Young, R. J. (2012). Behavioral responses of maned wolves (*Chrysocyon brachyurus*, Canidae) to different categories of environmental enrichment stimuli and their implications for successful reintroduction. *Zoo Biology*, 31, 453-469.

Davey, G. (2007). Visitors' effects on the welfare of animals in the zoo: a review. *Journal of Applied Animal Welfare Science*, 10, 169–183.

Fernandez, E. J., Tamborski, M. A., Pickens, S. R., & Timberlake, W. (2009). Animal-visitor interactions in the modern zoo: Conflicts and interventions. *Applied Animal Behavior Science*, 120, 1–8.

Fouraker, M., & Wagener, T. (Eds) (1996). AZA rhinoceros husbandry manual. Fort Worth, Texus: Fort Worth Zoological Park.

Goodenough, A. E., McDonald, K., Moody, K. & Wheeler, C. (2019). Are "visitor effects" overestimated? Behavior in captive lemurs is mainly driven by co-variation with time and weather. *Journal of Zoo and Aquarium Research*, 7, 59-66.

Gosling, S. D. (2001) From mice to men: what can we learn about personality from animal research? *Psychological Bulletin*, 127, 45–86.

Grand, A. P., Kuhar, C. W., Leighty, K. A., Bettinger, T. L., & Laudenslager, M. L. (2012). Using personality ratings and cortisol to characterize individual differences in African Elephants (*Loxodonta africana*). *Applied Animal Behavior Science*, 142, 69-75.

Grandia, P. A., van Dijk, J. J., & Koene, P. (2001). Stimulating natural behavior in captive bears. *Ursus*, 12, 199-202.

Hosey, G. (2000). Zoo animals and their human audiences: What is the visitor effect? Animal Welfare, 9, 343-357.

Höttges, N., Hjelm, M., Hård, T., & Laska, M. (2019). How does feeding regime affect behavior and activity in captive African lions (*Panthera le*o)? *Journal of Zoo and Aquarium Research*, 7, 117-125.

Hutchins, M., & Kreger, M. D. (2006). Rhinoceros behavior: implications for captive management and conservation. *International Zoo Yearbook*, 40, 150-173.

King, T., Boyen, E., & Muilerman, S. (2003). Variation in reliability of measuring behaviours of reintroduced orphan gorillas. *International Zoo News*, 50, 288-296.

Kuhar, C. W. (2008). Group differences in captive gorillas' reaction to large crowds. *Applied Animal Behavior Science*, 110, 377–385.

Kuhar, C. W. (2006). In the deep end: pooling data and other statistical challenges of zoo and aquarium research. *Zoo Biology*, 25, 339-352.

Larsen, M. J., Sherwen, S. L., & Rault, J. L. (2014). Number of nearby visitors and noise level affect vigilance in captive koalas. *Applied Animal Behavior Science*, 154, 76-82.

Lehner, P. N. (1998). Handbook of Ethological Methods. Cambridge: Cambridge University Press.

Less, E. H., Kuhar, C. W., Dennis, P. M., & Lukas, K. E. (2012). Assessing inactivity in zoo gorillas using keeper ratings and behavioral data. *Applied Animal Behavior Science*, 137, 74-79.

Maia, C. M., Volpato, G. L., & Santos, E. F. (2012). A case study: the effect of visitors on two captive pumas with respect to the time of the day. *Journal of Applied Animal Welfare Science*, 15, 222–235.

Margulis, S. W., Hoyos, C., & Anderson, M. (2003). Effect of felid activity on zoo visitor interest. *Zoo Biology*, 22, 587–599.

Melfi, V. A. (2009). There are big gaps in our knowledge, and thus approach, to zoo animal welfare: a case for evidence-based zoo animal management. *Zoo Biology*, 28, 574- 588.

O'Connor, S. M. (1986). Activity cycles of the southern white rhinoceros in captivity: implications for management. *International Zoo Yearbook*, 24/25, 297–303.

Orban, D. A., Siegford, J. M., & Snider, R. J. (2016). Effects of guest feeding programs on captive giraffe behavior. *Zoo Biology*, 35, 157-166.

Owen-Smith, N. (1973). The behavioral ecology of the white rhinoceros. Unpublished PhD thesis, University of Wisconsin, Madison, USA.

Penny, S. G., White, R. L., Scott, D. M., MacTavish, L., & Pernetta, A. P. (2019). Using drones and sirens to elicit avoidance behavior in white rhinoceros as an anti-poaching tactic. Proceedings of the Royal Society B, 286, p.20191135.

Penny, S. G., White, R. L., MacTavish, L., Scott, D. M., & Pernetta, A. P. (2020). Negligible hormonal response following dehorning in free-ranging white rhinoceros (*Ceratotherium simum*). *Conservation Physiology*, 8, coaa117.

Penny, S. G., White, R. L., Scott, D. M., MacTavish, L., & Pernetta, A.P. (2021). No evidence that horn trimming affects white rhinoceros horn use during comfort behavior and resource access. *Animal Biology*, 1, 1-17.

Quirke, T., O'Riordan, R. M., & Zuur, A. (2012). Factors influencing the prevalence of stereotypical behavior in captive cheetahs (*Acinonyx jubatus*). *Applied Animal Behavior Science*, 142, 189-197.

Rees, P. A. (2004). Low environmental temperature causes an increase in stereotypic behavior in captive Asian elephants (*Elephas maximus*). *Journal of Thermal Biology*, 29, 37-43.

Rose, P., & Riley, L. (2019). The use of Qualitative Behavioral Assessment to zoo welfare measurement and animal husbandry change. *Journal of Zoo and Aquarium Research*, 7, 150-161.

Ross, S. R. (2006). Issues of choice and control in the behavior of a pair of captive polar bears (*Ursus maritimus*). *Behavioral Processes*, 73, 117-120.

Sherwen, S. L., & Hemsworth, P. H. (2019). The visitor effect on zoo animals: implications and opportunities for zoo animal welfare. *Animals*, 9, 366-393.

Sherwen, S. L., Magrath, M. J., Butler, K. L., & Hemsworth, P. H. (2015). "Little penguins, *Eudyptula minor*, show increased avoidance, aggression and vigilance in response to zoo visitors". *Applied Animal Behavior Science*, 168, 71-76.

Vaz, J., Narayan, E. J., Kumar, R. D., Thenmozhi, K., Thiyagesan, K., & Baskaran, N. (2017). Prevalence and determinants of stereotypic behaviors and physiological stress among tigers and leopards in Indian zoos. *PloS One*, 12, 1-27.

Versteege, L. (2018). *EAZA White rhino EEP Best Practice Guidelines*. Safaripark Beekse Bergen: European Association of Zoos and Aquaria.

Whitham, J. C., & Wielebnowski, N. (2013). "New directions for zoo animal welfare science". *Applied Animal Behavior Science*, 147, 247-260.

Williams, R. L., Porter, S. K., Hart, A. G., & Goodenough, A. E. (2012). The accuracy of behavioral data collected by visitors in a zoo environment: can visitors collect meaningful data?. *International Journal of Zoology*, Article ID 724835.

Wood, W. (1998). Interactions among environmental enrichment, viewing crowds, and zoo chimpanzees (*Pan troglodytes*). *Zoo Biology*, 17, 211-230.

Yasui, S., Konno, A., Tanaka, M., Idani, G. I., Ludwig, A., Lieckfeldt, D., & Inoue-Murayama, M. (2013). Personality assessment and its association with genetic factors in captive Asian and African elephants. *Zoo Biology*, 32, 70-78.

Young, T., Finegan, E., & Brown, R. (2012). Effects of summer microclimates on behavior of lions and tigers in zoos. *International Journal of Biometeorology*, 57, 381-390.

Table 1: Ethogram for field recording, together with decision on retaining or grouping rarely-recorded behaviors before data analysis.

Category	Specific behavior recorded in field (n=23)	Data analysis decision (n=13)
Locomotion	Walking (moving at a steady pace)	Retained
	Running (moving quickly, not pursuing another rhino)	Retained
Inactive	Standing (usually head-down unless alert/vigilant)	Retained
	Resting - lying down / sleeping	Retained
Eating	Grazing (while walking slowly)	Retained
	Eating hay (sedentary)	Retained
	Suckling (juvenile drinking milk from mother)	Retained
Socializing	Social interaction between rhino; non-aggressive	Retained
Aggression	Horn jab (horn lowered, thrust towards another rhino)	} Two behaviors both uncommon; merged to create "Horn aggression"
	Horn swipe (sideways movement of horn on ground, displaying assertion of presence and/or status)	
	Chasing (pursuing another rhino)	Never witnessed; removed
Sensory	Sniffing urine or faces (sometimes with flehmen)	Singe observation; placed in "other"
	Vocalizing	Rarely witnessed; placed in "other"
Maintenance	Horn rubbing (against inanimate object)	Retained
	Urinating	} Two behaviors both uncommon so merged to create "Excretion"
	Defecating	
	Drinking	Rarely witnessed; placed in "other"
	Wallowing	Retained
	Body rubbing	Rarely witnessed; placed in "other"
Reproductive	Head resting (male head on female body)	} Three behaviors all uncommon; merged to create "Reproductive"
	Mounting (male climbs onto female trying to mate)	
	Mating (male mounts female and succeeds in mating)	
	Mouthing of genitals	Never witnessed; removed
Other	Other	Behaviors not observed frequently enough for a specific category – sum of sniffing, vocalizing, drinking and body rubbing
Unobservable	Out of sight	

Table 2: Researcher ethogram data on captive rhino compared to weather (ordinal scale running from 1 = heavy rain to 5 = sunny), temperature (ratio scale; higher = hotter), and visitor vehicles (ratio scale; higher = more). For overall rhino activity budget as summarized by PC1 from a Principal Components Analysis, a General Linear Model was used, which returned F values. For the frequency occurrence of seven most common individual behaviors, Generalized Linear Models with binomial error distribution and a logit link function were used, which returned χ^2 values).

	Weather			Temperature			Visitor vehicles		
	F/ χ^2	Dir	P	F/ χ^2	Dir	P	F/ χ^2	Dir	P
PC1	17.081	N/A	**<0.001**	71.842	N/A	**<0.001**	0.791	N/A	0.374
Grazing	146.884	-	**<0.001**	640.153	-	**<0.001**	0.805		0.370
Standing	18.529	-	**<0.001**	2.120		0.145	103.643	+	**<0.001**
Walking	5.050	-	**0.011**	42.741	-	**<0.001**	0.401		0.527
Eating hay	80.138	+	**<0.001**	103.087	+	**<0.001**	2.035		0.152
Resting	1.287		0.257	154.484	+	**<0.001**	295.807	-	**<0.001**
Running	0.002		0.962	1.329		0.249	0.262		0.609
Social interaction	1.636		0.201	3.242		0.072	2.098		0.147

Figure legends

Figure 1: Activity budget of all rhino across data collection period using researcher data and keeper data (Rub = Rubbing body, Sniff = Sniffing, Drink = Drinking, Voc = Vocalizing).

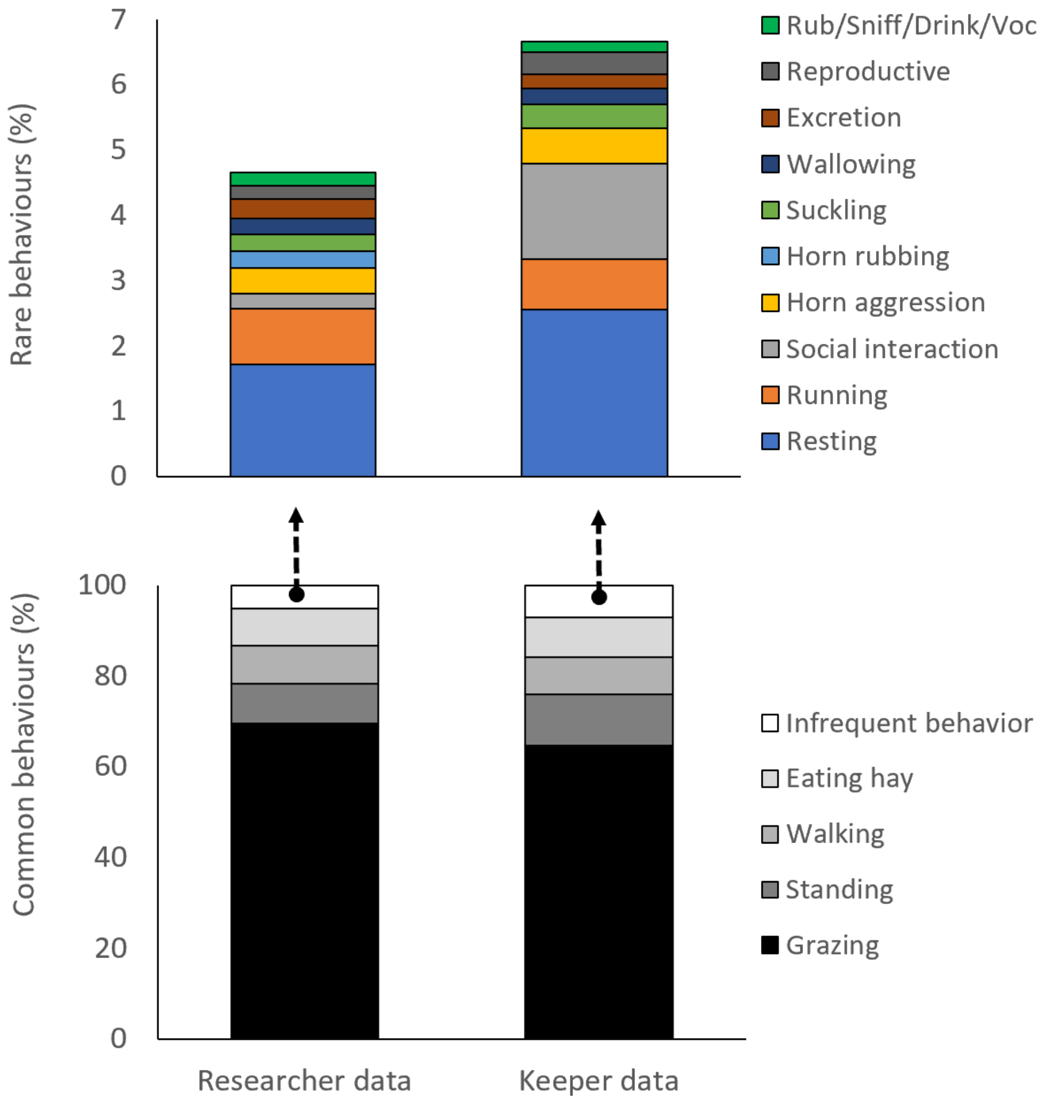

Figure 2: Frequency of occurrence of each behavioral category in pooled data from researchers (shaded) and keepers (unshaded). Data have been pooled across all rhino over the entire study period to allow comparison of datasets. Bars show mean per rhino per day; error bars show standard error.