Goodenough, Anne E ORCID logoORCID: https://orcid.org/0000-0002-7662-6670, Sewell, Amy and McDonald, Katie (2023) Behavioural patterns in zoo-housed Humboldt penguins (Spheniscus humboldti) revealed using long-term keeper-collected data: validation of approaches and improved husbandry. Applied Animal Behaviour Science, 258. ART 105811. doi:10.1016/j.applanim.2022.105811

Official URL: http://doi.org/10.1016/j.applanim.2022.105811
DOI: http://dx.doi.org/10.1016/j.applanim.2022.105811
EPrint URI: https://eprints.glos.ac.uk/id/eprint/11908

Behavioural patterns in zoo-housed Humboldt penguins (*Spheniscus humboldti*) revealed using long-term keeper-collected data: validation of approaches and improved husbandry

Anne E. Goodenough[1*], Amy Sewell[2], and Katie McDonald[2]

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

2 = West Midland Safari Park, Bewdley, UK

* Corresponding author

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

Orcid = 0000-0002-7662-6670

Abstract

Given the increasing obligation to elevate animal welfare beyond minimum expectations, zoos need robust mechanisms to monitor physical activity and species-appropriate behaviours. This is not without challenge as animal behaviour can vary seasonally or be influenced by external factors such as weather and visitors. In theory, keepers who work with animals year-round are ideally situated to collect behavioural data. However, time pressure often means any data collection is minimal, ad-hoc, or skewed towards particular times of day. Here we compare infrequent low-intensity ad-hoc data on activity and behaviour of zoo-housed Humboldt penguins (*Spheniscus humboldti*) collected by keepers with simultaneous high-intensity systematic data collected by a researcher. Once out-of-sight records (more frequent in researcher data) were excluded from both datasets, we found excellent agreement between keeper and researcher data. After this validation step, we collected extensive multi-year keeper data (diurnal hourly scan samples; 262 days across a 5-year study period) to assess potential influence of time of day, time of year, visitors, and weather on penguin behaviour. There were pronounced seasonal patterns: aquatic active behaviours were highest in autumn (38.9% in October) compared to early spring and late summer (February = 16.7%; August = 19.4%). Time of day was also important: terrestrial inactive behaviours increased throughout the day and this pattern was significantly more pronounced in summer than winter. Aquatic active behaviours were more frequent during warmer/sunnier conditions compared to cooler/wetter conditions, while the reverse was true for terrestrial inactive behaviours. Birds were much less active on days when the site was closed to visitors. Overall, we demonstrate that where keeper-collected data are validated and found to be robust and representative, there are real opportunities to study long-term behavioural patterns that can quickly translate to tangible husbandry benefits and improved animal welfare.

Keywords

Activity budget, captivity, ethogram, seasonal patterns, visitor effects, weather

1. Introduction

The obligation for zoological collections to elevate animal welfare beyond minimum expectations is becoming increasingly widely appreciated (Mellor et al., 2015; Wolfensohn et al., 2018). Animals need to be kept in enclosures that meet their physical health requirements are where healthy activity levels and species-appropriate behaviour are facilitated (McPhee and

Carlstead, 2012; Miller et al. 2020). Enclosures also need to be designed with visitors in mind in order to maximise experience and the predisposition of visitors to learn from zoo exhibits (Coe, 1985; Learmonth et al., 2021).

Although behavioural data should not be used as the only source of information when evaluating animal welfare (Rose and Riley, 2019; Watters et al., 2021), it is important that zoos have a reliable method of recording activity budgets and assessing how they correlate with external variables (Brando and Buchanan-Smith, 2018; Wark et al., 2019). Behavioural data can then be used to inform optimal husbandry (Whitham and Wielebnowski, 2013; Höttges et al., 2019), design of suitable enclosures (Ross, 2006; Quirke et al., 2012) and creation of high-quality enrichment (Coelho et al., 2012; Vaz et al., 2017). Activity budgets can also form part of the evidence used to examine behavioural diversity, which is increasingly being considered as a measure of welfare as well as cognitive stimulation (Miller et al., 2020; Brereton and Fernandez, 2022).

A key challenge when undertaking behavioural research on captive animals is the amount of data required to make evidence-informed decisions (Melfi, 2009). Behaviour can be influenced by temporal factors such as time of day (Grandia et al., 2001; Maia et al., 2012) and season (Cuculescu-Santana et al., 2017; Posta et al., 2013; Brando and Buchanan-Smith, 2018). Data ideally thus need to be both detailed (fine-scale; numerous datapoints per day) and to cover an extended temporal period. This is especially important where behaviour is affected by circannual events such as breeding and moulting (Serrano and Miller, 2000; Portugal et al., 2010) or temporally-changing variables such as temperature and weather (Rees et al., 2004; Young et al., 2012; Goodenough et al., 2019). Despite its importance, however, the majority of zoo-based studies use data collected over a short intense period. There are surprisingly few examples of zoo behavioural studies that explicitly consider seasonality but rare examples include Posta et al. (2013) on African elephant (*Loxodonta africana*), Rose et al. (2018) on flamingos (Phoenicopteridae), and Fernandez et al. (2020) on grizzly bears (*Ursus arctos horribilis*).

In theory, keepers are ideally situated to collect long-term behavioural data as they engage with animals on a daily basis throughout the year. This provides opportunities for regular recording over prolonged periods to ensure that data are not skewed by seasonal differences (Less et al., 2012; Carlstead et al., 2019). However, keepers are not always trained in standardised ethological data collection methods. It is notable that while previous research has utilised keeper assessments of animal personality (e.g. Grand et al., 2012; Yasui et al., 2013), ethogram-based research is typically undertaken by researchers. More importantly, as keepers' focus is necessarily on husbandry, time pressures mean that keeper datasets are usually smaller than researcher datasets, with fewer datapoints and lower sample sizes (Kuhar, 2006;). There might also be biases due to infrequent or ad-hoc sampling, especially if behaviour co-varies with time of day. It is, therefore, vital to cross-validate infrequent or ad-hoc keeper data with systematic researcher data (Gosling, 2001; Carlstead, 2009; Less et al., 2012). In situations where keeper data are proven to be robust and representative, there are real opportunities to use this framework to study long-term behavioural patterns, which can quickly translate to tangible husbandry benefits and improved animal welfare (Goodenough et al., 2022).

Penguins, and especially Humboldt penguins (*Spheniscus humboldti*), are popular with visitors and commonly held by zoos (Stevenson et al., 1994; Hernandez-Colina, et al., 2021). The species is thus a good "test case" for studying whether keepers can collect robust and representative data. There are also practical husbandry reasons why recording long-term behavioural data for Humboldt penguins is particularly valuable. Swimming is a vital aspect of penguin behaviour: except during moult, wild penguins spend >80% of their time foraging at sea and sometimes remain continuously in the water for >24hr (Taylor et al. 2002; Eriacher-Reid et al., 2012; de la Puente et al., 2013). However, in captivity baseline aquatic activity levels are considerably lower, ranging from 2% to 23% in different collections (Marshall et al., 2016). A low level of swimming in captivity is regarded as indicative of poor physical and cognitive health (Clarke, 2003; Marshall et al., 2016). Moreover, this is concerning for animal welfare because captive penguins are prone to ulcerative pododermatitis (bumblefoot), a bacterial infection of the feet. Prolonged weight-bearing on terrestrial substrate and lack of activity have been suggested as possible contributors to bumblefoot in captive penguins (Clark & Kerry, 2000, Eriacher-Reid et al., 2012, Reisfeld et al., 2013). Low levels of aquatic behaviour are also concerning from a visitor perspective. It is important that penguins show species-appropriate aquatic behaviours so exhibits can be used to educate visitors about the threats of oceanic pollution and unsustainable resource (Collins et al., 2016, 2020). Moreover, studies on little penguin (*Eudyptula minor*) have shown increased activity, especially swimming, is associated with better visitor experience (Chiew et al., 2019, 2022). Previous research has suggested swimming might be influenced by external factors such as visitor presence (Sherwen et al., 2015, Collins et al., 2016), feeding method (Fernandez et al., 2021), and enclosure design (Marshall et al., 2016). However, patterns are not consistent between studies and most research does not capture natural seasonal variation (e.g. during the moult phase of the circannual rhythm). This is that identifying appropriate husbandry changes is challenging.

Our aims in this paper are: (1) to compare low-intensity infrequent keeper-collected data on activity and behaviour of Humboldt penguins with simultaneous high-intensity detailed researcher-collected data; (2) use an extensive multi-year keeper dataset (diurnal hourly scan samples for 262 days across a 5-year period) to create a robust activity budget; and (3) assess the impacts of time of day, time of year, visitor presence, and weather conditions on the amount of time birds devote to terrestrial inactive behaviours versus aquatic active behaviours. We discuss our results both from a methodological standpoint (i.e. testing keeper data collection frameworks) and in relation to husbandry improvements for Humboldt penguins as the focal species in this paper.

2. Methods

2.1 Study setup

This study was conducted at West Midland Safari Park (WMSP, Worcestershire, UK) between December 2015 and March 2020. The Humboldt penguin (*Spheniscus humboldti*) colony initially comprised 16 adults; numbers then fluctuated between 16 and 23 due to a small number of hatchings, acquisitions and deaths. The enclosure contained an internal space used at night, surrounded by a larger external space used during the day. Penguins were released from overnight accommodation into the external space ~09:00 and remained outside until at least 15:00 depending sunset and opening hours. The external part of the enclosure included a deep pool covering ~40% of the total surface area. Terrestrial areas were covered by three different substrates: render moulded to resemble bedrock (~35%), stones and pebbles overlying render (~15%) and sand overlying render (~10%). Wooden and render nestboxes were located throughout the external exhibit. Visitor viewing was possible on three sides of the external exhibit, including two areas where glass formed part of the edge of the pool to underwater viewing. A schematic of the enclosure is provided as supplementary material.

2.2 Penguin data collection

Data were collected using an ethogram via passive observation when birds were in the outdoor exhibit. The ethogram listed 11 mutually-exclusive behaviours within three activity categories (terrestrial inactive, terrestrial active, aquatic active) plus “other” and “out-of-sight” (Table 1). When “other” was used, the behaviour was allocated *aposteriori* to the relevant activity category based on the description given. There were five people involved with data collection: one researcher (co-author KS, research officer for the focal collection) and a team of four keepers under the leadership of co-author AS. To reduce the risk of inter-observer variation, ethogram definitions were agreed by five data recorders. Before data collection formally started, training was undertaken whereby observers watched exactly the same birds at the same time, recorded behaviour independently, and then discussed their data to identify and understand differences. This was repeated until there was >95% agreement in pilot data, assessed using Cohan's Kappa as per Lehner (1996). This initial work was necessary to ensure that validation of the keeper data using researcher data was not confounded or masked by inter-observer variation.

2.3 Penguin datasets

Penguin behaviour data were collected at group level using instantaneous (pinpoint) scan sampling (Altmann, 1974). This involved recording the number of birds undertaking each behaviour at pre-determined "pinpoint" times, which recent computer simulation work has demonstrated to be a more accurate method for recording behaviour than using interval (one-zero) methods (Brereton et al., 2022). To minimise risk of the presence of observers confounding penguin behaviour (Hosey, 2008), data were collected from outside the enclosure with observers standing as far distanced as possible while retaining a clear view of the enclosure. Two datasets were collected:

- Keeper data: collected using one instantaneous scan sample per hour from 09:00 to 15:00, as close as possible to the top of the hour. This gave (up to) 7 datapoints per day with each datapoint summarising the behaviour of the entire group of penguins (16-23 individuals). To ensure data were representative of general behaviour rather than temporary changes during feeding, data were not collected during the two short scheduled public penguin feeds. Keeper data were collected on 262 days across the 5-year study period (1,580 scan samples; mean 6.03 samples per day). The sample effort distribution was: 2015 = 3 days (Dec only); 2016 = 85 days; 2017 = 51 days; 2018 = 49 days; 2019 = 52 days; 2020 = 22 days (Jan-Mar only before COVID-19 lockdowns prematurely ended data collection). The overall dataset was a composite of data collected by four keepers.

- Researcher data: collected six times per day using instantaneous scan sampling at 2-minute intervals for a 30-minute time period. Normally this gave 16 datapoints per time period, but occasionally recording was slightly truncated due to public feeding events starting early (to match the keeper dataset, researcher data were never collected during scheduled public feeds). Each of the six 30-minute data time periods started as close as possible to the top of the hour thus: 10:00-10:30, 11:00-11:30, 12:00-12:30, 13:00-13:30, 14:00-14:30, 15:00-15:30. The final period was only feasible when penguins were out at least 15:30 and was missed on some short winter days. When all data collection periods were feasible and no data collection period was truncated, there were 96 datapoints per day with each datapoint summarising the behaviour of 16-23 penguins. Researcher data were collected on 13 days across the 5-year period, all directly matching keeper data collection days. This dataset was compiled by a single researcher.

The 13 days when both keeper and researcher data were collected comprise is henceforth termed the comparison dataset; the full (multi-year) dataset contained keeper data only. Researcher data were condensed by summing all observations in each recording period for each behaviour to give one line of data containing percentage frequency of each behaviour over that time period. This not only avoided pseudo-replication (i.e. multiple lines of non-independent data within the same recording period) but also ensured that these data aligned with the single record per hour in the keeper data, thereby allowing the uneven datasets to be directly compared.

2.4 Visitor presence and weather data collection

Visitor presence was a binary variable determined by whether the Safari Park was open to the public (216 days) or closed to the public (46 days). Hourly records of temperature (°C), precipitation (mm), and windspeed (km/h) were available from a weather station within 1 km (latitude 52.3886°N, longitude 2.2497°W). Datapoints for the period January 2016 to March

2020 were downloaded from http://visualcrossing.com/weather (n = 37,201) and each penguin record was matched to the relevant weather record using a unique year/month/day/hour (YYYYMMDDHH) reference in both datasets. Weather data were not available for the three recording days in December 2015.

2.5 Data analysis

To establish whether the low-intensity data collected by keepers matched high-intensity researcher data (such that the former could be considered an accurate reflection of penguin behaviour), one activity budget was created using researcher data and a second was created using the comparison data collected by keepers on the same days. For this comparison, the first keeper datapoint at ~09:00 was excluded so the overall period when data were recorded for a given day matched between keeper and researcher datasets (research data collection commenced at 10:00). Statistical analysis was undertaken using Z Tests for Proportions to compare frequency occurrence of each activity category between the datasets, which allowed differences in individual behavioural categories to be compared statistically. Analysis was initially run on the raw data and then on recalculated percentage data after removal of out-of-sight records from both datasets (i.e. to rescale each activity budget to include only records where penguins had been sighted so behaviour could be recorded to ensure that differential occurrence of out-of-sight did not skew proportional occurrence of other behaviours).

After the initial keeper-research verification analysis, we used the full multi-year keeper-collected dataset (262 days; 09:00-15:00) to create a full activity budget. To get a baseline understanding of the data, we graphically related the frequency occurrence of terrestrial inactive behaviours and frequency occurrence of aquatic active behaviours to external factors of: (1) month of year; (2) time of day; and (3) visitor presence as a binary variable based on whether the Safari Park was open or closed to visitors. As month of year was a circular variable, such that December and January are adjacent categories, circular histograms were created as per Batschelet (1981) and Mardia and Jupp (1999) using Oriana Circular Statistics (Version 4, Kovach Computing Services).

Then, to explore data in more detail and replicate the real-world complexity in which multiple external factors co-occur, a Generalised Linear Model (GLM) framework was used. Two models were created, one for behaviours in the terrestrial inactive category and one for behaviours in the aquatic active category. To allow for there being different number of penguins in the group at different times, the dependent variable was the number of penguins undertaking the focal activity in relation to the number of penguins observed at that datapoint. A binomial error distribution and a logit link function was used. Three fixed factors were entered as predictors: (1) month; (2) time of day; and (3) whether the Safari Park was open/closed to visitors and interactions were created thus: month * time of day; open/closed * month; and open/closed * time of day. Inclusion of these interactions substantially improved the fit of both models (i.e. decreased delta Akaike's Information Criterion scores by >10: Burnham and Anderson, 2002; Burnham et al., 2011). Four continuous covariate predictors were added: (1) temperature (min = -2.4°C; max = 29.9°C); (2) precipitation (min = 0 mm/hr; max = 6.21 mm/hr), (3) cloud cover (min = 0%; max = 100%); and (4) windspeed (min = 1.3 km/h; max = 39.1 km/hr). Entering month as a fixed factor (i.e. categorical variable) rather than a covariate meant that no underlying linear distribution was assumed for this inherently circular variable. Models were run using SPSS version 27 and used data from 259 days when weather data were available, giving a sample size of 1,562 hourly observations (18 records over three days in December 2015 were excluded).

2.6 Ethics

In this study, the focal penguins remained in their normal enclosure with no changes to routine husbandry, no handling of animals, and no manipulation of the captive environment.

Behavioural data were collected by their assigned keepers from outside the enclosure without any form of interaction. No changes were relative to standard keeping practice for the birds.

3. Results

3.1 Validation of keeper data

Initial analysis suggested that there were significant differences in the frequency of two of the four activity categories (Table 2). However, further investigation revealed that this was entirely due to penguins being out-of-sight being more frequently in the researcher data (6.8%) compared to the keeper data (2.8%). Because of the mutually-exclusive method used to record behaviour, this skewed the relative frequency occurrence of behaviours in other activity categories. In other words, the fact that keepers were less likely to use the out-of-sight category relative to the researcher caused a fundamental mismatch in the activity budgets. When out-of-sight records were removed from both datasets, so that just the data where penguins had been observed were used, there were no significant differences (Table 2).

A follow-up analysis comparing individual behaviours between researcher and keeper comparison datasets showed that there was excellent agreement (Figure 1). The only minor differences, in opposing directions, were for standing (keeper = 37.3%; researcher = 40.7%) and resting (keeper = 12.0%; researcher = 9.7%), but discussion with keepers indicated these differences were not large enough to result in them forming different conclusions or make different husbandry recommendations, especially as they were in the same activity category. This comparative analysis thus provided good evidence that keeper-collected data were accurate and reliable and could be used for the rest of the project across multiple years to generate robust data over a prolonged time period.

3.2 Penguin behavioural patterns: baseline analysis

The most common behaviours exhibited by the penguin group over the full 262-day study period between 2016 and 2020 (based on keeper data with out-of-sight excluded) were: standing (46.4%), swimming (19.8%), resting (14.6%), terrestrial auto-preening (8.9%), aquatic auto-preening (4.5%), and terrestrial locomotion (3.9%). The remaining behaviours were only witnessed rarely (<0.5%). It is notable that swimming occurred less often, and standing occurred more often, in the full keeper-collected dataset than in the keeper-collected subset used for the comparison with researcher data (swimming 19.8% vs 33.8% in Fig 1; standing 46.4% vs 40.4% in Fig 1). These differences are likely due to the larger sample size and being more comprehensive in spanning seasons and different weather conditions in the full multi-year dataset, which itself underlines the need to collect data over multiple seasons to fully understand behaviour.

As one of the two motivations for this study, alongside comparison of keeper and research datasets, was to understand the factors influencing the amount of time penguins spend inactive on land compared to being active in the water, subsequent analysis focused on terrestrial inactive behaviours and aquatic active behaviours. Terrestrial inactive behaviours – combining standing and resting – accounted for 60.7% of the penguin activity budget on average using the percentage of observations. There was variation across the year (monthly lowest = 43.3% in October; monthly highest 69.8% in February; Figure 2a). Aquatic active behaviours – combining swimming, aquatic auto-preening, and aquatic feeding – accounted for 24.6.% of the activity budget on average using the percentage of observations. There was an autumn peak (38.9% in October) contrasting with much lower rates in early spring (16.7% in February) and late summer (19.4% in August) (Figure 2b). The frequency occurrence of behaviours also varied with time of day (Figure 3), although patterns were less pronounced than for time of year. Generally terrestrial inactive behaviours increased throughout the day from a low of 52.9% at 09:00 to a high of 68.3% at 15:00, while aquatic active behaviours varied from 17.0% to 29.3%, generally peaking at 12:00 and 14:00 (prior to the scheduled public feeding events, which were not themselves recorded).

Whether the Safari Park was open or closed, and therefore whether there were visitors on site or not, also affected behaviour. The average frequency occurrence of terrestrial inactive behaviours in November-February was 67.1% on closed days decreasing to 56.8% for open days. The average frequency of aquatic active behaviours in this same period showed the opposite pattern, rising from 19.7% on closed days to 27.2% on open days. The difference in terrestrial inactive and aquatic inactive in relation to open/closed status was especially pronounced in November (Figure 4).

3.3 Penguin behavioural patterns: multivariate modelling

The multivariate GLM models demonstrated that terrestrial inactive behaviours and aquatic active behaviours were both influenced by external factors. As expected given graphical output (Figs 2-4), month of year, time of day, and whether the Safari Park was open to visitors were all statistically significant predictors of the amount of time penguins spent undertaking terrestrial inactive and aquatic active behaviours (Table 3). There were also interactions between these variables. Particularly notable was that time of day patterns were more pronounced: (1) during summer compared to winter; and (2) during periods when the Safari Park was open compared to when it was closed (Table 3). Weather also had an effect, with aquatic active behaviours occurring significantly more frequently in warmer/sunnier conditions compared to cooler/wetter conditions when inactive terrestrial behaviours were more common (Table 3). Precipitation and wind speed were non-significant in both models.

4. Discussion

4.1 Validation of keeper data

Our results indicate that, for this group of Humboldt penguins, there was excellent agreement between high-intensity data collected systematically by a dedicated researcher and low-intensity data collected by keepers. Activity budgets were very similar and there were no statistical differences between datasets for activity categories, but only once out-of-sight records had been excluded to remove the effect of the out-of-sight category being used more often in the researcher data. The statistical differences in relative frequency of out-of-sight records (higher in researcher data) might be due to keepers being unwilling to record out-of-sight if they were concerned this would reflect badly on their professionalism (e.g. Goodenough et al., 2022). Alternatively, keepers might have made inferences about what out-of-sight birds were likely doing. The fact that keepers were only collecting a single instantaneous scan sample rather than multiple instantaneous scan samples in quick succession might also have allowed keepers more opportunity to locate more birds. We recommend that keepers be asked to locate as many individuals as possible, but also to use out-of-sight as necessary to ensure data collection is accurate. This is especially important if out-of-site individuals are likely to be engaging in specific behaviours (e.g. using nestboxes to rest: Marshall et al. (2016)). Although it is possible to remove out-of-sight records post-hoc and rescale data as we have done here, avoiding the problem initially is preferable.

Throughout this study, care was taken to minimise any influence of observers on the penguins being studied (Hosey, 2008). Animals can respond differently to humans according to whether they are keepers, non-keeper staff, or visitors (e.g. Mitchell et al., 1991; Melfi and Thomas, 2005), and can be modified by other factors (e.g. Thompson (1989) found ungulates displayed heightened vigilance and alert responses to keepers when the zoo was closed). This further emphasises the need to verify keeper data with non-keeper or video-capture data before keeper data are relied upon. Once keeper data reliability has been checked, the potential for keepers to go beyond qualitative observations or categorical assessments of activity and personality (e.g. Grand et al., 2012; Less et al., 2012; Yasui et al., 2013) to monitor behaviour quantitively is exciting. Keeper-led quantitative assessment has considerable potential to directly inform husbandry, as demonstrated by Carlstead (2009) for several species including cheetah (*Acinonyx jubatus*) and maned wolf (*Chrysocyon brachyurus*). It is also a valuable tool within research contexts, as shown by work on gorilla (*Gorilla gorilla*) activity in relation to age/sex groupings (Less et al., 2012) and the effect of temperature, weather conditions, and visitor numbers on behaviour of white rhino (*Ceratotherium simum*) (Goodenough et al., 2022).

For the benefits of keepers collecting animal behaviour data to be realised on the ground rather than remaining hypothetical possibilities, it is vital that data are easy to collect and that trends can be reviewed and visualised "real time". To facilitate this at the focal collection, we developed a system to enable keepers to use tablets or phones to enter data directly into a pre-programmed spreadsheet. This prevents reliance on pen and paper and negates the need for a data entry step. The spreadsheet is set up so that it auto-generates monthly profiles of behaviour in real time and the output is visible to keepers, veterinary staff, and researchers. The precepts are similar to those underpinning the flexible and customisable Zoo Monitor platform (Wark et al., 2019) but using a general software spreadsheet package rather than an App interface. Use of this system supported the anecdotal view that the penguin group was spending a large part of their activity budget standing. This informed an enclosure redesign with substantially more sand and rocks (and less render) to minimise the impact of prolonged periods of weightbearing on hard terrestrial substrates on foot health (Clark & Kerry, 2000, Eriacher-Reid et al., 2012, Reisfeld et al., 2013). The opportunities for such initiatives to be applied much more widely for a range of species (subject to an initial keeper data verification step being completed satisfactorily) are substantial. The approach offers genuine potential for improving the behaviour and welfare of captive animals in ways that are not only cost- and time-efficient, but that also unite the research process and keeper experience in ways that are inclusive and mutually beneficial.

4.2 Penguin behavioural patterns

Many behavioural studies in zoos are undertaken over relative short time periods (weeks or months). This might be because the time constraints on keepers are such that data collection is hard to fit into daily schedules, especially over a long period, or sometimes because research often relies upon short-term student placements or internships. Activity budgets in published literature are thus often temporally limited (e.g. Simeone et al., 2002; studies included in meta-analysis of Marshall et al., 2016). This means that any seasonal patterns (Serrano and Miller, 2000; Portiagl et al., 2010; Brando and Buchanan-Smith, 2018) or variation relative to weather variables (e.g. Rees et al., 2004; Young et al., 2012; Goodenough et al., 2019) are often disregarded.

Our findings show that the studied penguins spend comparatively little time undertaking behaviours classed as aquatic active (mean = 24.6%). Although higher than reported at some other collections (range = 2-23%; mean = 9.8%; Marshall et al., 2016) this is still low compared to the behavioural profile in the wild (Simeone et al., 2002; Eriacher-Reid et al., 2012; de la Puente et al., 2013). On its own, this type of baseline information would be of limited use for improving enclosure design or informing husbandry changes. However, the low-intensity keeper-collected framework adopted here allowed long-term data collection that revealed the drivers for behavioural patterns, which does become useful in informing on-ground change to improve animal welfare. Patterns in aquatic activity (and, conversely, terrestrial inactivity) are likely partly related to seasonal breeding and moulting, which influence both captive (Marshall et al., 2016) and wild (Taylor et al., 2002; de la Puente et al., 2013) penguins. However, external factors, such as visitor presence and husbandry factors (e.g. feed times) were important mediating factors.

Studying factors that influence behaviour graphically (and univariately) is useful to get a baseline understanding of the data to share with keepers. However, multivariate models can elucidate the significance of patterns and disentangle how co-occurring factors interact with one another in more complex ways. Here, low levels of aquatic activity for this group of penguins during the winter had been hypothesised based on casual observations. Multivariate modelling revealed that aquatic active behaviours are actually higher in November and December than in many other months on average, but that aquatic activity is very low on days when the Safari Park is closed to visitors, which occurs only in the winter. This example of where external factors are superimposed upon, and interact with, intrinsic circadian patterns is likely due to visitor presence (and increased pool-based feeding during scheduled public feeds) encouraging penguins into the water. In the wild, penguins' main motivation for entering the water is feeding (de la Puente et al., 2013). Hand feeding in captivity means tha/t the motivation to spend time in the water is reduced (Clarke, 2003, Marshall et al., 2016). Although terrestrial hand feeding can be necessary to ensure appropriate food share and administer medication as per Taxon Advisory Groups Best Practice Guidelines (EAZA 2005; AZA 2014), we concur with Marshall et al. (2016) that pool-feeding is a potentially effective way to increase aquatic activity and this could involve live fish to further increase swimming before feed times (Fernandez et al., 2021). We recommend that this pool-based feeding is used wherever possible for at least part of the daily food provision, and that this be done year round (i.e. that becomes the default husbandry rather than something undertaken primarily to improve visitor experience).

4.3 Conclusions and recommendations

This long-term project provides useful information for other penguin holders, by generating robust evidence that different levels of aquatic activity at different times of year and times of day should be expected in captive penguin colonies. Future studies that look at behaviour only over one season, no matter how well designed or data rich, need to take this into account. More generally, we have demonstrated that keeper-collected data can provide a time-efficient, long-term dataset that gives an ‘expected’ baseline pattern of variation in activity over seasons and time of day. This can be used to confirm anecdotal observations, reveal patterns of behaviours not immediately apparent, and allow rapid identification of changes that could be indicative of a potential problem.

Acknowledgements

We thank Head and Deputy Head of Wildlife at West Midland Safari Park, Angela Potter and Noel Carey, for their support throughout the project, penguin keepers Vicky McFarlane, Abbey Stone and Holly Palmer-Wilson for help with data collection, and Rachel Wallace for assistance with data entry. We also acknowledge four students from the University of Gloucestershire, Anna Jones, Louise Chiverton, Cheryl Greaves and Megan Warren, whose dissertation projects acted as a catalyst for this research.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-of-profit sectors.

References

Altmann, J. (1974). Observational study of behavior: Sampling methods. *Behavior* 49, 227–266.

AZA. (2014). *Taxon Advisory Group: Penguin (*Spheniscidae*) Care Manual*. Maryland: Association of Zoos and Aquariums.

Batschelet, E (1981). *Circular Statistics in Biology*. London: Academic Press.

Brereton, J. E., & Fernandez, E. J. (2022). Investigating Unused Tools for the Animal Behavioral Diversity Toolkit. *Animals*, 12, 2984.

Brereton, J. E., Tuke, J., & Fernandez, E. J. (2022). A simulated comparison of behavioural observation sampling methods. *Scientific Reports*, 12, 1-10.

Brando, S. & Buchanan-Smith, H. M. (2018). The 24/7 approach to promoting optimal welfare for captive wild animals. *Behavioural Processes*, 156, 83-95.

Burnham, K. P., Anderson, D. R. (2002). *Model Selection and Multimodel Inference: A Practical Information-Theoretical Approach*. 2d ed. New York: Springer-Verlag.

Burnham, K. P., Anderson, D. R., and Huyvaert, K. P. (2011). AIC model selection and multimodel inference in behavioral ecology: some background, observations, and comparisons. *Behavioral Ecology and Sociobiology*, 65, 23-35.

Carlstead, K. (2009). A comparative approach to the study of keeper–animal relationships in the zoo. *Zoo Biology*, 28, 589-608.

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

Chiew, S. J., Hemsworth, P. H., Melfi, V., Sherwen, S. L., Burns, A., & Coleman, G. J. (2021). Visitor attitudes toward little penguins (*Eudyptula minor*) at two Australian zoos. *Frontiers in Psychology*, 12, 626185.

Chiew, S. J., Hemsworth, P. H., Sherwen, S. L., Melfi, V., & Coleman, G. J. (2019). The effect of regulating zoo visitor-penguin interactions on zoo visitor attitudes. *Frontiers in Psychology*, 10, 2351.

Clark, J. & Kerry, K. (2000). Diseases and parasites of penguins. *Penguin Conservation.* 13, 5-24.

Clarke, A G. (2003). Factors affecting pool use by captive Humboldt penguins (*Spheniscus humboldti*). Proceeding of the 5th annual symposium of zoo research, Marwell Zoological Park, UK, pp 190–204.

Coe, J. C. (1985). Design and perception: Making the zoo experience real. *Zoo Biology*, 4, 197-208.

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.

Collins, C. K., Quirke, T., Overy, L., Flannery, K., & O'Riordan, R. (2016). The effect of the zoo setting on the behavioural diversity of captive gentoo penguins and the implications for their educational potential. *Journal of Zoo and Aquarium Research*, 4, 85-90.

Collins, C., Corkery, I., McKeown, S., McSweeney, L., Flannery, K., Kennedy, D., & O'Riordan, R. (2020). An educational intervention maximizes children's learning during a zoo or aquarium visit. *The Journal of Environmental Education*, 51, 361-380.

Cuculescu-Santana, M., Horn, C., Briggs, R.N., Bowe, C., & Geraughty, M.L. (2017). Seasonal changes in the behaviour and enclosure use of captive Asian small clawed otters. *IUCN Otter Specialist Interest Group Bulletin* 34, 29 – 50.

de la Puente, S., Bussalleu, A., Cardeña, M., Valdés-Velasquez, A., Majluf, P., & Simeone, A. (2013). Humboldt penguin (*Spheniscus humboldti*), in: Garcia-Borboroglu, P., & Boersma, P. D. (Eds.). *Penguins: Natural History and Conservation*. Seattle: University of Washington Press, pp. 265–283.

Erlacher-Reid, C., Dunn, L., Camp, T., Macha, L., Mazzaro, L. & Tuttle, A. D. (2012). Evaluation of potential variables contributing to the development and duration of plantar lesions in a population of aquarium-maintained African penguins (*Spheniscus demersus*). *Zoo Biology*, 31, 291–305.

EAZA (2005). *Spheniscus* penguin husbandry manual. EAZA Penguin Taxon Advisory Group.

Fernandez, E. J., Myers, M., Hawkes, N. C. (2021). The effects of live feeding on swimming activity and exhibit use in zoo Humboldt penguins (*Spheniscus humboldti*). *Journal of Zoological and Botanical Gardens*, 2, 88–100

Fernandez, E. J., Yoakum, E., & Andrews, N. (2020). Seasonal and daily activity of two zoo-housed grizzly bears (*Ursus arctos horribilis*). *Journal of Zoological and Botanical Gardens*, 1, 1-12.

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

Goodenough, A. E., Price, T. W., Brazier, S. L., & McDonald, K. (2022). Factors affecting the behavior of captive white rhinoceros (*Ceratotherium simum*) and the accuracy of ad-hoc keeper data. *Zoo Biology*. 1–10. https://doi.org/10.1002/zoo.21723.

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 Behaviour Science*, 142, 69-75.

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

Hernandez-Colina, A., Gonzalez-Olvera, M., Eckley, L., Lopez, J., & Baylis, M. (2021). Avian malaria affecting penguins in zoological gardens, aquariums and wildlife parks in the UK. *Vet Record*, 189, e511.

Hosey, G. (2008). A preliminary model of human–animal relationships in the zoo. *Applied Animal Behaviour Science*, 109, 105-127.

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

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

Learmonth, M. J., Chiew, S. J., Godinez, A., & Fernandez, E. J. (2021). Animal-visitor interactions and the visitor experience: Visitor behaviors, attitudes, perceptions, and learning in the modern zoo. *Animal Behavior and Cognition*, 8, 632-649.

Lehner, P. N. (1996). 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 Behaviour 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.

Mardia, K.V. & Jupp, P.E. (2000). Statistics of directional data. London: John Wiley and Sons.

Marshall, A. R., Deere, N. J., Little, H. A., Snipp, R., Goulder, J., & Mayer-Clarke, S. (2016). Husbandry and enclosure influences on penguin behavior and conservation breeding. *Zoo Biology*, 35, 385-397.

McPhee, M. E. & Carlstead, K. (2012) The importance of maintaining natural behaviors in captive mammals, in: Kleiman, D. G., Thompson, K. V. & Kirk Baer, C. (Eds.) Wild Mammals in Captivity: Principles and Techniques for Zoo Management. Chicago, University of Chicago Press, pp 303-313.

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.

Melfi, V. A., & Thomas, S. (2005). Can training zoo-housed primates compromise their conservation? A case study using Abyssinian colobus monkeys (*Colobus guereza*). *Anthrozoös*, 18, 304-317.

Mellor, D. J., Hunt, S. & Gusset, M. (Eds.) (2015). Caring for Wildlife. The World Zoo & Aquarium Animal Welfare Strategy. Gland: WAZA Executive Office, 87 pp.

Miller, L. J., Vicino, G. A., Sheftel, J., & Lauderdale, L. K. (2020). Behavioural diversity as a potential indicator of positive animal welfare. *Animals*, 10, ref: 1211.

Mitchell, G., Obradovich, S. D., Herring, F. H., Dowd, B., & Tromborg, C. (1991). Threats to observers, keepers, visitors, and others by zoo mangabeys (*Cercocebus galeritus chrysogaster*). *Primates*, 32, 515-522.

Portugal, S. J., Isaac, R., Quinton, K. L., & Reynolds, S. J. (2010). Do captive waterfowl alter their behaviour patterns during their flightless period of moult? *Journal of Ornithology*, 151, 443-448.

Posta, B., Huber, R., & Moore, D. E. (2013). The effects of housing on zoo elephant behavior: a quantitative case study of diurnal and seasonal variation. *International Journal of Comparative Psychology*, 26, 37-52.

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

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

Reisfeld, L., Barbirato, M., Ippolito, L., Cardoso, R. C., Nichi, M., Sgai, M. G., & Pizzutto, C. S. (2013). Reducing bumblefoot lesions in a group of captive Magellanic penguins (*Spheniscus magellanicus*) with the use of environmental enrichment. *Pesquisa Veterinária Brasileira*, *33*, 791-795.

Rose, P. E., Brereton, J. E., & Croft, D. P. (2018). Measuring welfare in captive flamingos: Activity patterns and exhibit usage in zoo-housed birds. *Applied Animal Behaviour Science,* 205, 115-125.

Rose, P.E., & Riley, L. (2019). The use of Qualitative Behavioural 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 behaviour of a pair of captive polar bears (*Ursus maritimus*). *Behavioural Processes*, 73, 117-120.

Serrano, A., & Miller, E. H. (2000). How vocal are harp seals (*Pagophilus groenlandicus*)? A captive study of seasonal and diel patterns. *Aquatic Mammals*, *26*, 253-259.

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

Stevenson, M. F., Dryden, H., Alabaster, A., & Wren, C. (1994). The new penguin enclosure at Edinburgh Zoo: the palace for the 1990s. *International Zoo Yearbook,* 33, 9-15.

Taylor, S. S., Leonard, M. L., Boness, D. J., & Majluf, P. (2002). Foraging by Humboldt penguins (*Spheniscus humboldti*) during the chick-rearing period: general patterns, sex differences, and recommendations to reduce incidental catches in fishing nets. *Canadian Journal of Zoology,* 80, 700-707.

Thompson, V.D., 1989. Behavioral response of 12 ungulate species in captivity to the presence of humans. *Zoo Biology*, 8, 275-297.

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

Wark, J. D., Cronin, K. A., Niemann, T., Shender, M. A., Horrigan, A., Kao, A., & Ross, M. R. (2019). Monitoring the Behaviour and habitat use of animals to enhance welfare using the ZooMonitor app. *Animal Behaviour and Cognition*, 6, 158-167.

Watters, J. V., Krebs, B. L., & Eschmann, C. L. (2021). Assessing Animal Welfare with Behavior: Onward with Caution. *Journal of Zoological and Botanical Gardens,* 2, 75–87.

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

Wolfensohn, S., Shotton, J., Bowley, H., Davis, S.,Thompson, and Justice, W. S. M. (2018). Assessment of Welfare in Zoo Animals: Towards Optimum Quality of Life. *Animals,* 8, ref 110.

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 Biometeorolog*y, 57, 381-390.

Table 1: Ethogram behaviours used within this study. As data were not recorded during scheduled public penguin feeds, asterisked behaviours relating to feeding were rarely observed: "terrestrial feeding" was never used; "aquatic feeding" was retained because penguins sometimes retrieved fish from the bottom of the pool outside formal feeding times.

Activity Category	**Behaviour**	**Definition**
Terrestrial Inactive	Standing	Stationary in an upright position; weight bearing on feet
	Resting	Stationary in a horizontal position; weight bearing on body
Terrestrial Active	Terrestrial locomotion	Locomotion on land in an upright position, including short hops
	Terrestrial auto-preening	Contact between bird's bill and own feathers while on land
	Terrestrial allo-preening	Contact between bird's bill and feathers of another bird while on land when this is not thought to be part of courtship
	Courtship	Initiating/receiving courtship behaviours: pursuit, head rubbing, calling, embrace, mounting, energetic flipper movement
	Nest building	Active interest in, and manipulation of, material used for nesting: collection from exhibit or placement in nestbox
	*Terrestrial feeding	Consuming fish or actively looking to obtain food while on land
Aquatic Active	Swimming	Movement in water, other than preening or feeding, including both surface and below surface swimming, and porpoising
	Aquatic auto-preening	Contact between bird's bill and own feathers while in water
	*Aquatic feeding	Consuming fish or actively looking to obtain food while in water
Other	Other	Give full description and allocate to Terrestrial Inactive, Terrestrial Active, or Aquatic Active
Out-of-sight	Out-of-sight	Individual cannot be observed

Table 2: Comparison of activity categories between keeper-collected data and researcher-collected data using original data and then recalculated after exclusion of out-of-sight records.

	Original data				Recalculated data (excluding out-of-sight)			
	Keeper (%)	Researcher (%)	Z	P	Keeper (%)	Researcher (%)	Z	P
Terrestrial inactive	48.3	46.7	1.092	0.276	49.6	50.1	-0.313	0.757
Terrestrial active	10.0	10.4	-0.296	0.764	10.4	11.2	-0576	0.562
Aquatic active	38.8	36.1	1.065	**0.049**	39.9	38.7	0.840	0.401
Out-of-sight	2.8	6.8	-5.244	**<0.001**				

Table 3 Generalised Linear Models examining the influence of temporal, visitor and weather variables on frequency on terrestrial inactive behaviours (left) and aquatic active behaviours (right). A binomial distribution with a logit link function was used in both models and model fit was considered using Akaike's Information Criterion when adding interaction terms to ensure that they improved the model, such that model fit versus parsimony was optimised.

Independent variable		d.f.	Terrestrial Inactive Behaviours (Standing & Resting)			Aquatic Active Behaviours (Swimming, Aquatic Auto-Preening & Aquatic Feeding)		
			Wald χ^2	P	Explanation	Wald χ^2	P	Explanation
Factors	Month of year	11	599.750	**<0.001**	Highest in spring (Fig 2a)	383.928	**<0.001**	Highest in autumn (Fig 2b)
	Time (hr 09:00-15:00)	6	34.456	**<0.001**	Highest late in day (Fig 3)	72.146	**<0.001**	Highest middle of day (Fig 3)
	Park open or closed	1	162.446	**<0.001**	Higher when closed (Fig 4)	130.769	**<0.001**	Higher when open to visitors (Fig 4)
Covariates	Temperature	1	10.536	**0.001**	Higher when colder	6.353	**0.012**	Higher when warmer
	Precipitation	1	1.112	0.292		2.885	0.089	

	Cloud cover	1	0.334	0.564		4.233	**0.040**	Higher when sunny
	Wind speed	1	2.783	0.095		0.010	0.919	
Interactions	Month * time of day	66	335.851	<**0.001**	Time of day patterns more pronounced in summer; less pronounced in winter	389.126	<**0.001**	Time of day patterns more pronounced in summer; less pronounced in winter
	Open/closed * month	3	46.086	<**0.001**	Always higher when closed but magnitude of effect varies between months (Fig. 4)	69.274	<**0.001**	Always higher when open but magnitude of effect varies between months (Fig. 4)
	Open/closed * time	6	35.185	<**0.001**	Uniformly high when closed to visitors, increasing throughout the day when open to visitors	27.136	<**0.001**	Uniformly low when closed to visitors, peaks middle day when open to visitors

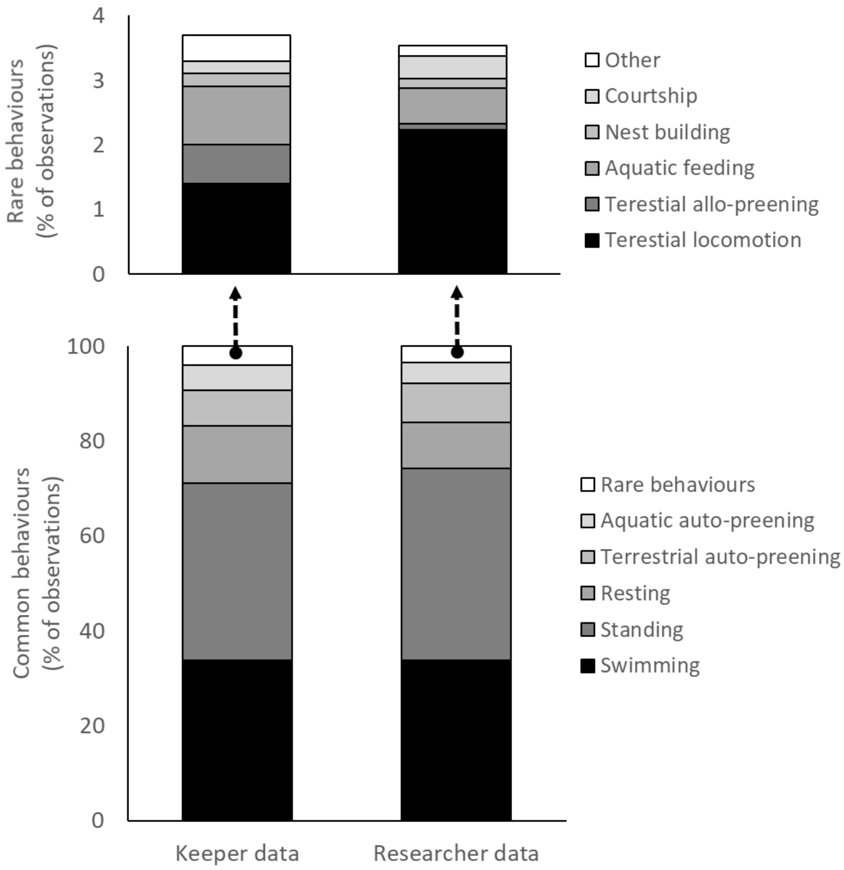

Figure 1: Activity budget of all penguins using keeper data and researcher data from the 13 days (96 data recording periods) when both types of data were collected simultaneously. For ease of display, specific behaviours where the number of observations was <2% of the total observations were grouped together under the name "rare behaviours" on the main (lower) histogram, with the exact percentage occurrence of each specific behaviour in this category shown graphically in the expanded (top) histogram.

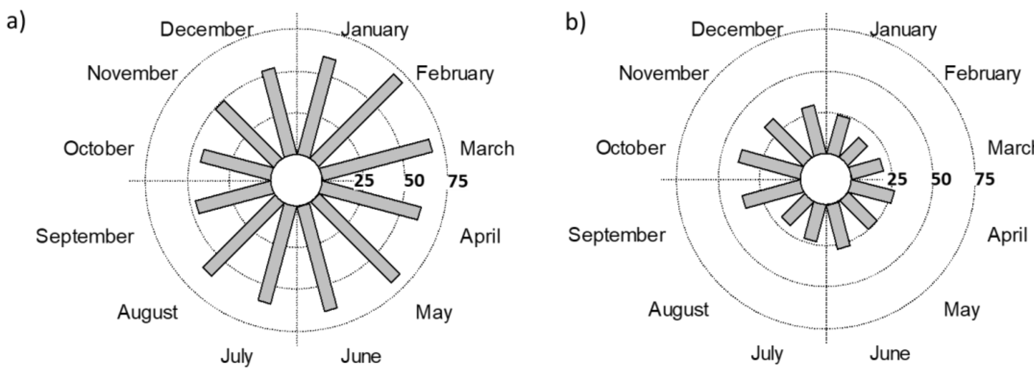

Figure 2: Percentage occurrence of: (a) terrestrial inactive behaviours related to month of year; and (b) aquatic active behaviours relative to month of the year. In both cases, bars

represent the mean occurrence of that behavioural category in the penguin activity budget per month and the concentric gridlines represent 25%, 50% and 75%, from the centre of each circular histogram to its outer edge. Graphs were drawn using Oriana Circular Statistics (Version 4, Kovach Computing Services).

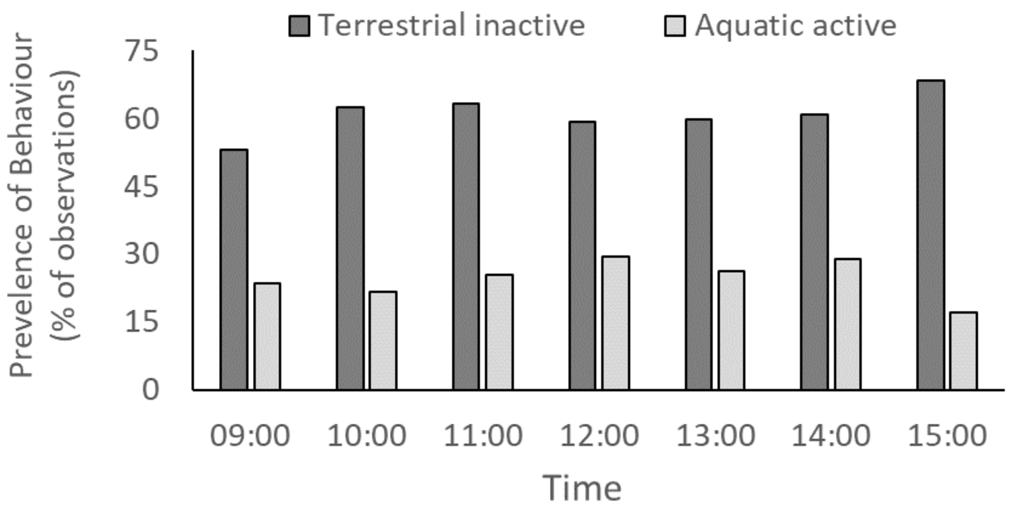

Figure 3: Percentage occurrence of terrestrial inactive behaviours (dark bars; combination of standing and resting) and aquatic active behaviours (light bars; combination of swimming, aquatic auto-preening and aquatic feeding) in relation to time of day.

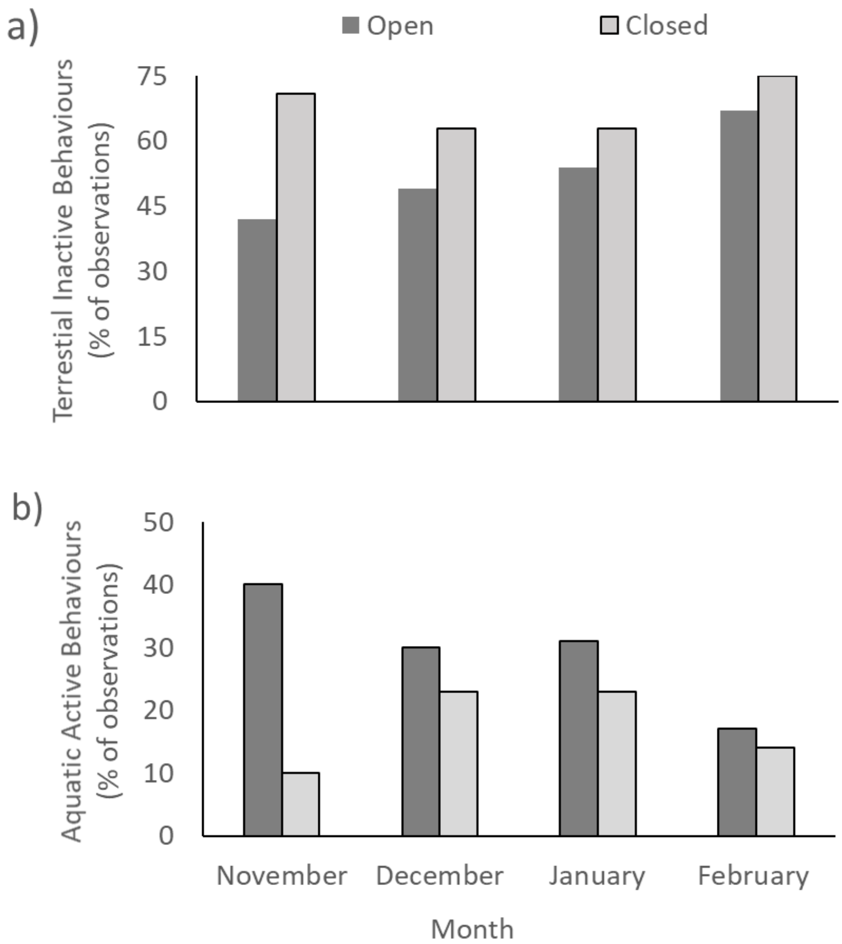

Figure 4: Percentage occurrence of behaviour in relation to when the safari park was open to visitors (dark bars) or closed to visitors (light bars): (a) terrestrial inactive behaviours (combining standing and resting); and (b) aquatic active behaviours (combining swimming, aquatic auto-preening and aquatic feeding).

1. Declaration of interests

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

☐The authors declare the following financial interests/personal relationships which may be considered as potential competing interests:

Highlights

- Ad-hoc keeper data on penguin behaviour validated using systematic researcher data
- Once verified, informative multi-year keeper data collected with minimal effort
- Important aquatic active behaviour lowest in spring, cold weather & when zoo closed
- At focal collection understanding behavioural influences informed enclosure changes
- Verified keeper data are valuable for recording behaviour (general applicability)