

Behavioural diversity as a potential welfare indicator for professionally managed chimpanzees (*Pan troglodytes*): Exploring variations in calculating diversity using species-specific behaviours

K Hall^{*†‡}, J Bryant[‡], M Staley^{‡§}, JC Whitham[‡] and LJ Miller[‡]

[†] Sedgwick County Zoo, 5555 W Zoo Boulevard, Wichita, KS 67212, USA

[‡] Chicago Zoological Society – Brookfield Zoo, 3300 Golf Road, Brookfield, IL 60513, USA

[§] Department of Biology, Loyola University Chicago, Chicago, IL 60660-1537, USA

* Contact for correspondence: KatieHallPhD@gmail.com

Abstract

Behavioural diversity may serve as a positive indicator of animal welfare that can be applied in long-term monitoring schemes in managed settings (eg zoos, laboratories, farms). Behavioural diversity is often higher when animals live in stimulating environments and experience positive events. Unfortunately, welfare researchers have not adopted consistent, standardised approaches to measuring behavioural diversity. The goal of this exploratory study was to utilise data from 41 adult chimpanzees (*Pan troglodytes*) housed across 16 zoological institutions to examine various models of Shannon's Diversity Index. Specifically, we investigated the impact of: combining versus splitting behaviours, including only positive behaviours, including human interaction, and considering recipient behaviours. We evaluate how the inclusion or exclusion of different behaviours impacts the relationship of behavioural diversity with: (i) concentrations of faecal glucocorticoid metabolites (GCM), a common indicator of adrenal activity; (ii) concentrations of immunoglobulin-A (IgA), an indicator of immune function and potential indicator of positive welfare; and (iii) stereotypic behaviour, a validated indicator of poor welfare. Most indices had significant negative relationships with faecal GCM. Animals that express a variety of behaviours from their species-typical repertoire have lower average faecal GCM concentrations and are likely experiencing better welfare. We did not find significant relationships between the behavioural diversity indices and IgA concentrations. Two indices were inversely associated with stereotypic behaviour. Our findings provide additional support for using Shannon's Diversity Index to calculate behavioural diversity as a robust, valid measure of positive welfare. However, future publications must justify the process for including or excluding behaviours from calculations.

Keywords: animal welfare, behavioural diversity, chimpanzee, faecal glucocorticoid metabolites, immunoglobulin-A, Shannon diversity index

Introduction

As animal welfare research continues to incorporate indicators of good welfare into long-term monitoring schemes, and specifically animal-based measures that provide insight into an individual's physical, mental, and emotional states (Butterworth *et al* 2011; Siegford 2013; Whitham & Wielebnowski 2013), behavioural diversity may serve as a positive indicator that can be applied in managed settings (eg zoos, laboratories, farms, sanctuaries, shelters). After all, when individual animals in managed settings express a diverse repertoire of species-typical behaviour, they are presumed to have good welfare (Rushen *et al* 1993; Shepherdson *et al* 1993; Wemelsfelder *et al* 2000; Swaisgood *et al* 2001, 2005; Swaisgood & Shepherdson 2006; Miller *et al* 2011, 2016; but see Vickery & Mason 2004). However, very few welfare studies measure behavioural diversity, and of those that do, there is inconsistency in their methodology. Our exploratory study utilises data from professionally managed chimpanzees (*Pan troglodytes*) to examine various models of behavioural diversity.

There are two indices that can be employed to integrate behavioural diversity into welfare research. Behavioural diversity is most often calculated using a formula borrowed from ecology, the Shannon Diversity Index (H) (Shannon & Weaver 1949). Whereas in ecology, H measures the proportion of individuals belonging to each of a number of species, ethologists and welfare researchers have adapted this index to measure the proportion of time an individual spends in particular behavioural categories, and values are affected by both the number of categories (abundance) and the distribution of values among them (evenness) (for a review, see Miller *et al* 2020). However, when behaviour is distributed evenly among categories, it is more appropriate to use Simpson's Diversity Index (D) (Simpson 1949). Simpson's Diversity Index, again deriving from ecology, gives more weight to dominant or common behaviours and is less affected by an animal spending a short amount of time in a few additional behavioural categories. For instance, an individual who distributes time equally across ten behavioural

categories should be measured using D, and an individual who distributes time unevenly across ten categories should be measured using H. When behavioural diversity is higher, the value of D is lower, but H is higher. Specifically, when additional behaviours are included in the calculation of Shannon's H, even though there are more behaviours, each occurs at a lower frequency and therefore does not significantly impact the value of the index, which takes into account both number and frequency of behaviour. For example, if an H diversity index was calculated using a general category of 'play' and another index was calculated using both 'solitary play' and 'social play', despite an additional behaviour in the latter index, the final values would only be slightly higher due to the lower frequencies within each behavioural category.

Though very few studies measure behavioural diversity, and even fewer apply diversity indices such as Shannon's H or Simpson's D, the theoretical lynchpin is that the measure should include common species-typical behaviours, while excluding inactivity (unless it can be distinguished from lethargy) and any abnormal, stereotypic or repetitive behaviours (Swaigood *et al* 2001; Miller *et al* 2011; Razal *et al* 2016). The argument for utilising behavioural diversity as an indicator of welfare is based on the idea that when the index is high, there is an increased likelihood that we are meeting an animal's behavioural needs and when the index is low, the animal is more likely to be stereotyping or lethargic, both potential signs of compromised welfare (Mason & Latham 2004; Meagher & Mason 2012; Meagher *et al* 2017). Diversity of behaviour, by definition, suggests that an individual distributes his/her time among a variety of behaviours. Stereotypic behaviour is typically not included in calculations of behavioural diversity because it is often associated with behavioural restriction. Individuals who are behaviourally restricted may be motivated to perform a behaviour but are unable to do so due to the current environment and conditions (Würbel *et al* 1996; Mason *et al* 2001; Sarrafchi & Blokhuis 2013). Indeed, engaging in stereotypy at high levels (eg pacing), not only reflects lower behavioural diversity (eg sows [*Sus scrofa*]: Stolba *et al* 1983; laboratory rabbits [*Oryctolagus cuniculus*]: Gunn & Morton 1995; but see Vickery & Mason 2004 for Asiatic black bears [*Ursus thibetanus*] and Malayan sun bears [*Helarctos malayanus*]) but also is typically associated with reduced welfare (Duncan & Petherick 1991; Fraser & Duncan 1998). For example, sows living in restricted stall conditions engaged in only 33 behaviours, many stereotyped, compared to sows housed in a semi-natural environment, who engaged in over 100 behaviours (Stolba *et al* 1983). The fact that the sows in stalls engaged in fewer behaviours overall and spent more time stereotyping suggests that their restrictive environment likely resulted in a narrower behavioural repertoire. Inactive behaviour — unless it can be clearly distinguished from lethargy — should also be excluded from behavioural diversity calculations, as lethargy could be a sign that behavioural needs are not being met (Meagher & Mason 2012; Meagher *et al* 2017). Finally, different durations of inactivity are expected for various

species (eg consider lions [*Leo panthera*] and hummingbirds [*Trochilidae*]), which would complicate the interpretation of index comparisons.

Behavioural diversity is often found to be higher when animals live in stimulating environments or experience positive events. In many cases, providing animals with an enriched environment results in greater behavioural diversity, diminishes stereotypic behaviour, and promotes other behaviours indicative of good welfare (Swaigood *et al* 2001). For example, after gaining access to a climbing structure and new enrichment items, spectacled bears (*Tremarctos ornatus*) exhibited a decrease in stereotypic behaviours and an increase in behavioural diversity (Renner & Lussier 2002). Similarly, pandas (*Ailuropoda melanoleuca*) displayed a reduction in stereotypic behaviour, greater behavioural diversity, and increased activity levels when presented with enrichment items (Swaigood *et al* 2001). This suggests that providing individuals with various challenges, stimulating enrichment, novel environmental features and the opportunity to make choices can promote behavioural diversity.

One challenge of comparing behavioural diversity within species is the inconsistency of methodologies (Cronin & Ross 2019), which is why we focus our discussion only on Shannon's H (the Simpson index is less useful in ethology because it is unlikely that behaviour will be evenly distributed among categories). Renner and Lussier (2002) discuss diversity of behaviour without reporting the use of a particular index. Wemelsfelder and colleagues (2000) acknowledge that more behavioural opportunities were available to pigs housed in substrate-enriched conditions compared to those living in a substrate-impooverished environment, and therefore calculated relative diversity by controlling for the maximum diversity possible for each condition. Pigs in the enriched condition had greater relative behavioural diversity than those in the impoverished condition, when observed in their home pens and while participating in novel object experiments. Furthermore, the pigs living in substrate-impooverished conditions were more likely to exhibit other behaviours indicative of compromised welfare, including higher levels of fear behaviour and reduced mobility (ie less walking). Rather than calculate diversity based on relative opportunities available to different groups, it would have been more straightforward to compare values of the same index. Indeed, a standardised, widely accepted index would allow researchers to make comparisons across individuals, conditions, and locations.

While we see the value in applying Shannon's H as a standardised metric, we also understand there are questions that must be addressed. Cronin and Ross (2019) thoroughly discussed four main concerns with this index: (i) how the size of the behavioural repertoire may influence H's responsiveness; (ii) how the resolution of the behavioural categories (ie combining data with modifiers into the overarching behaviour vs splitting the behaviours with modifiers) may impact H's sensitivity; (iii) that the valence of the behaviours is not considered in the calculations; and (iv) that some behavioural categories are commonly excluded from calculations. The current study attempts to address some of these concerns.

Specifically, the goal of the current study was to collect data on professionally managed chimpanzees to examine behavioural diversity, and specifically Shannon's Diversity Index, as a welfare indicator. Furthermore, we explore whether differences exist when including or excluding different behavioural categories and compare these indices to physiological and behavioural indicators of animal welfare. We first examine traditional models for including behaviours in diversity indices (ie incorporating a wide range of species-typical positive behaviours performed by the focal animal). As so few studies in the fields of welfare science and ethology have used this metric, we also felt it necessary to investigate whether differences arise from the addition of less-traditional behaviours. This included examining models with both positive (eg grooming) and negative (eg contact aggression) behaviours as well as human interaction. Chimpanzees, like many species under managed care, experience both positive and negative behaviours, and may interact with staff and zoo visitors on a regular basis. Finally, we also consider recipient behaviours, because being the recipient of an action, such as grooming or aggression may impact welfare differently than being the agent of the same action.

As physical and psychological stress can influence glucocorticoid concentrations (Sapolsky 2002) and immunoglobulin A (IgA) concentrations (Tsujita & Morimoto 1999; Bishop & Gleeson 2009), we evaluate how the inclusion or exclusion of different behaviours in an index calculation impacts the relationship of behavioural diversity with faecal glucocorticoid metabolites (GCM) and IgA. Specifically, faecal GCM is a common indicator of adrenal activity and an individual's physiological response to perceived or actual threats (ie stressors) (Moberg & Mench 2000). Although some stressors are considered beneficial and part of an animal's natural history (eg courtship, hunting, arousal), consistently high levels of faecal GCM — caused by exposure to chronic or repeated, acute stressors — may be associated with the dysregulation of metabolic (Elsasser *et al* 2000), reproductive (Moberg 1987), and immune systems (Rivier 1995; Blecha 2000). IgA is an indicator of immune function and a potential indicator of positive welfare (for a review, see Staley *et al* 2018). There is evidence that IgA increases after both humans and non-human animals experience positive emotional states or pleasant stimuli (Pressman & Cohen 2005; Watanuki & Kim 2005). For example, both salivary and faecal IgA have been utilised as a biomarker of stress in rats (*Rattus norvegicus*) (Guhad & Hau 1996; Eriksson *et al* 2004). For dogs (*Canis familiaris*), Skandakumar and colleagues (1995) reported that salivary IgA was negatively correlated with salivary cortisol and associated with behavioural assessments. Specifically, low levels of IgA were associated with scores indicative of 'a dog exhibiting stress' and high levels with dogs who were considered 'calm' and 'confident' (Skandakumar *et al* 1995). In a series of studies on shelter cats, researchers found that cats (*Felis silvestris catus*) who received a petting treatment

(four times daily for ten minutes over a period of ten days) had higher faecal IgA, were more likely to maintain their mood, and were less likely to exhibit behaviours associated with anxiety and frustration (Gourkow *et al* 2014; Gourkow & Phillips 2015). Additionally, cats participating in positive reinforcement training had higher faecal IgA compared to controls (Gourkow & Phillips 2016).

By comparing behavioural diversity with physiological and behavioural metrics, we aim to provide additional data on behavioural diversity as a positive welfare indicator. We predicted that higher behavioural diversity would be associated with lower faecal GCM and higher IgA concentrations. We also predicted that behavioural diversity would be associated with spending less time exhibiting stereotypic behaviours. We expected that the Shannon Diversity Index would be a robust indicator of welfare, despite manipulations such as adding behavioural categories to the calculation.

Materials and methods

Study approval

Approval to conduct this research was given by animal care and veterinary staff at each participating institution. The three months of data that were analysed for this study were part of a larger project. No manipulations were made to the subjects' care or husbandry routine during these three months of data collection.

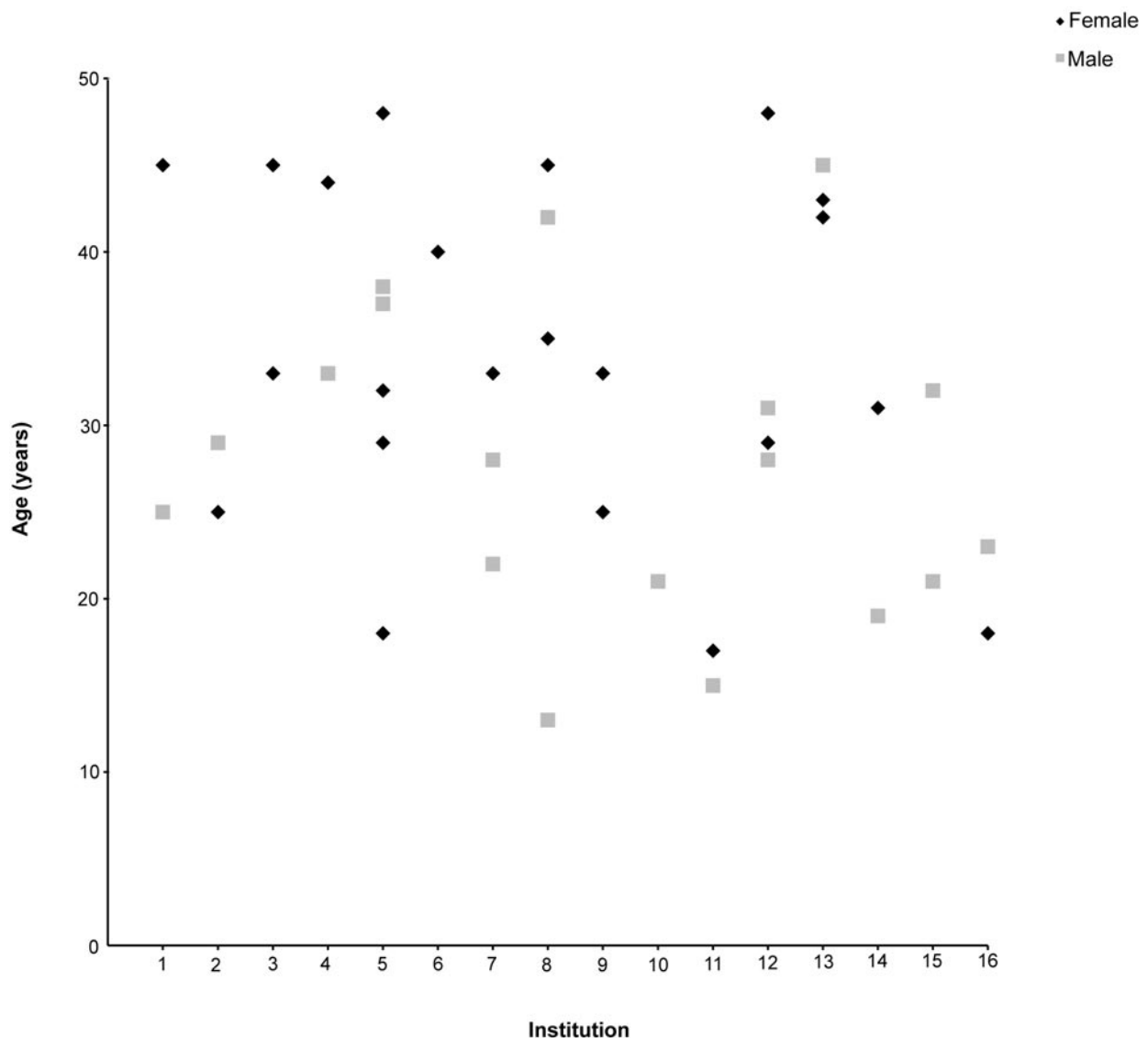
Study animals

Subjects included 41 adult chimpanzees (18 males, 23 females) with no known health concerns at the start of the study (Figure 1). Subjects were recruited from 16 zoological institutions accredited by the Association of Zoos and Aquariums. Each participating institution indicated how many chimpanzees in their collection could contribute data to the study and provided a list of all individuals. After excluding sub-adults and individuals with known health concerns, subjects were chosen using a random number table. Each subject contributed behavioural and physiological data for three months.

Behavioural observations

Volunteers or staff from each institution filmed each subject for a 30-min focal follow observation, three times per week between 1 March and 28 May, 2016. Videos were filmed during alternate mornings and afternoons. Videos were coded by staff and trained volunteers from the Chicago Zoological Society (CZS) using BORIS software (Torino, Italy; Friard & Gamba 2016) according to a specified ethogram adapted from Ross and Lukas (2001). The behavioural items analysed for this study and their definitions are presented in Table 1 (see supplementary material to papers published in *Animal Welfare*: <https://www.ufaw.org.uk/the-ufaw-journal/supplementary-material>). All staff and volunteers reached $r > 0.80$ inter-rater reliability before coding videos. Behaviours were scored as mutually exclusive continuous state variables on a focal individual.

Figure 1



Subject demographics.

Physiological measures

Animal care professionals at each institution aimed to collect daily faecal samples from each subject between 1 March and 28 May, 2016. Specifically, they were asked to collect the first defaecation of the day, to reduce intra-individual differences that may have arisen due to circadian rhythms (Heintz *et al* 2011; Staley *et al* 2018) and to make inter-individual comparisons more reliable. Subjects housed with conspecifics received baking-grade food colouring (AmeriColor™ soft gel paste, Placentia, CA, USA) to allow care professionals to distinguish between the study animals' samples. Samples were stored in a -20°C freezer to maintain stability of analytes until they could be shipped to CZS for hormone and IgA analyses. All samples were then shipped to CZS overnight on dry ice and upon receipt stored in a -20°C

freezer. A total of 3,226 samples were collected from 41 individuals in this study (mean: 79 samples per individual).

Using an analytical scale (Mettler balance, model #AB104-5, Mettler, OH, USA), $0.5 (\pm 0.05)$ g of each faecal sample was weighed into two separate 16×125 mm (diameter \times height) polypropylene tubes for faecal GCM and IgA analysis. Samples were processed using wet weights, as wet weight and dry weight have been found to correlate moderately well (Palme *et al* 2013). Each tube was labelled with sample number, animal ID, date, and whether it was designated for faecal GCM or IgA analysis. Exact weights were recorded on a datasheet and subsequently entered into a Microsoft Excel® spreadsheet. Approximately 5 g of leftover faeces from each sample was placed into in a 12×75 mm polypropylene tube as a backup. All tubes were then stored at -20°C until use.

Faecal glucocorticoid metabolite (GCM) measurement

One day prior to hormone analyses, faecal GCM tubes were removed from the -20°C freezer and extracted using 5 ml of 80% ethanol in dH_2O . Tubes were vortexed for approximately 30 s and then placed overnight on a rotator set to 30 rotations per min (Fisher Labline Maxi Rotator, model #4631, Fisher Scientific, Waltham, MA, USA). The next morning, tubes were centrifuged at 1,500 rpm for 15 min (Marathon 3000R centrifuge, model #120, Fisher Scientific). One milliliter of supernatant from each sample was pipetted into a new 12×75 mm polypropylene tube containing 1 ml of assay buffer (0.1 M phosphate buffered saline containing 1% BSA, pH 7.0) to produce a 1:10 dilution. Samples were assayed immediately following the extraction using a commercially available corticosterone EIA kit (catalogue #901-097, Enzo Life Sciences, Ann Arbor, MI, USA). All plates were read on a spectrophotometer (Dynex MRX Revelation, Dynex, Chantilly, VA, USA) at an optical density of 405 nm.

The techniques used for faecal GCM analysis have been used in previous studies (eg Chinnadurai *et al* 2009; Bashaw *et al* 2016), and a biological validation (ie an adrenocorticotrophic hormone, or ACTH, challenge) for chimpanzees has been conducted (saliva: Heintz *et al* 2011; faeces: Murray *et al* 2013). Variability between assays (inter-assay coefficient of variation [CV]) was monitored using high and low controls across all plates. To determine intra-assay variability, a single sample was repeated ten times on a single plate.

Biochemical validation of the faecal GCM assay consisted of a linearity test to determine parallelism with the standard curve, in addition to a recovery test to measure the concentration of exogenous analytes. To establish parallelism, serial two-fold dilutions of a sample pool were tested for potential interference in the sample matrix, linearity with the standard curve, and to determine the appropriate dilution factor at which to run the samples. The optimal sample dilution for faecal GCM was 1:500, as this dilution was closest to 50% binding of the sample pool.

Recovery of exogenous faecal GCM was measured by spiking one diluted sample with each of the five highest standards in separate tubes. Each standard contained a known amount of hormone ranging from 250–4,000 pg ml^{-1} . The average percent recovery was calculated by dividing the measured concentration of faecal GCM by the expected concentration of faecal GCM multiplied by 100.

The cross-reactivity of the Enzo Life Sciences corticosterone antibody are 100% corticosterone, 28.6% desoxycorticosterone, 1.7% progesterone, 0.28% tetrahydrocorticosterone, 0.18% aldosterone, 0.13% testosterone and any other steroids were $< 0.05\%$. Assay sensitivity was 26.99 pg ml^{-1} and the intra-assay coefficient of variation was 3.94% at 80.48% binding with an average concentration of 38.76 pg ml^{-1} ($n = 10$). Inter-assay variation was determined using a high and low control, 10.05% CV at 28.08% binding and 19.15% CV at 49.86% binding, respectively. The mean (\pm SD) average recovery of exogenous corticosterone was $103.34 (\pm 24.50)\%$. All faecal GCM concentrations were expressed as ng g^{-1} wet faeces.

Immunoglobulin A (IgA) measurement

Immunoglobulin A samples were extracted and assayed following the methods of Lantz *et al* (2016). Briefly, IgA was extracted using 5 ml of 1X phosphate buffered solution (PBS; 5.42 g NaH_2PO_4 , 8.66g Na_2HPO_4 , 8.7g NaCl, 0.8 g NaOH added to 1 L dH_2O , pH adjusted to 7.2 using 5M NaOH). Tubes were vortexed for 30 s, placed on the rotator (Fisher Labline Maxi Rotator, model #4631) for 2 h, and then centrifuged at 1,500 rpm for 15 min. One milliliter of supernatant was then transferred to a new 1.7 ml polypropylene Eppendorf tube and stored in a -20°C freezer until analysis. We assayed the samples using a commercially available IgA-human ELISA (catalogue #E80-102, Bethyl Laboratories, Montgomery, TX, USA). Plates were read using a spectrophotometer (Dynex MRX Revelation) at an optical density of 405 nm.

Biochemical validation of the IgA assay consisted of a linearity test to determine parallelism with the standard curve, in addition to a recovery test to measure the concentration of exogenous analytes. To establish parallelism, serial two-fold dilutions of a sample pool were tested for potential interference in the sample matrix, linearity with the standard curve, and to determine the appropriate dilution factor at which to run the samples. The optimal sample was 1:75 for IgA.

The Bethyl Laboratories IgA-human antibody is 100% specific to human IgA, and no further testing on cross-reactivity with other species has been studied at this time (Lantz *et al* 2016 physiologically validated the use of a human-specific enzyme immunoassay by performing an ACTH challenge in a chimpanzee). There is no cross-reactivity with other human immunoglobulins or serum proteins. Assay range is 7.8–500 ng ml^{-1} . The intra-assay coefficient of variation was 4.23% at 17.76% binding with an average concentration of 69.98 ng ml^{-1} ($n = 10$). Inter-assay variation was determined using a high and low control, 13.15% CV at 23.45% binding and 17.96% CV at 79.38% binding, respectively. We expressed all IgA concentrations in $\mu\text{g g}^{-1}$ wet faeces.

Statistical analysis

Data from the three-month data collection period were compiled in Microsoft Excel®. After correcting for time visible during the observation periods, we calculated the proportion of time each subject spent in each behavioural category. We calculated the Shannon Diversity Index using the formula:

$$H' = -\sum_{i=1}^R p_i \ln p_i$$

The Shannon Diversity Index is most appropriate given the structure of our data (ie time unevenly distributed among many behavioural categories) (Shannon & Weaver 1949). In this formula, p_i is the proportion of time belonging to the i th behavioural category in the dataset, multiplied by the natural logarithm of this proportion ($\ln p_i$); this product is summed across behaviours and multiplied by -1 . R represents the total number of behavioural categories included in the index of interest. For each index, p_i is recalculated as categories are included or removed. Due to the fact that the total number of behavioural categories varies across

Table 3 Descriptive statistics for each behavioural diversity index.

Measure	N	Min–Max H	Mean (\pm SD)	R	Max possible H	Min H as % of max possible H	Max H as % of max possible H	Mean H as % of max possible H
Index 1	41	1.11–2.07	1.70 (\pm 0.20)	11	2.40	46.25%	86.25%	70.83%
Index 2	41	0.75–1.61	1.29 (\pm 0.17)	6	1.79	41.90%	89.94%	72.07%
Index 3	41	1.17–2.11	1.79 (\pm 0.20)	17	2.83	41.34%	74.56%	63.25%
Index 4	41	0.78–1.69	1.37 (\pm 0.17)	10	2.30	33.91%	73.48%	59.57%
Index 5	41	1.27–2.16	1.85 (\pm 0.20)	20	3.00	42.33%	72.00%	61.67%
Index 6	41	0.84–1.75	1.44 (\pm 0.18)	11	2.40	35.00%	72.92%	60.00%
Index 7	41	1.21–2.18	1.87 (\pm 0.22)	21	3.04	39.80%	71.71%	61.51%
Index 8	41	0.74–1.68	1.35 (\pm 0.17)	10	2.30	32.17%	73.04%	58.70%

For the Shannon Diversity Index (H), the proportion of time spent in each behavioural category (p_i) sums to 1. For each index, p_i is recalculated as categories are included or removed. Due to the fact that the total number of behavioural categories (R) varies across indices, the maximum possible value of H also differs.

Table 4 Descriptive statistics for faecal glucocorticoid metabolites (GCM), immunoglobulin-A (IgA) and stereotypic behaviour.

Measure	N	Min–max	Mean (\pm SD)
GCM (ng g ⁻¹)	41	26.97–1,318.58	188.45 (\pm 282.39)
IgA (μ g g ⁻¹)	41	10.29–1,963.89	72.67 (\pm 303.08)
Stereotypic behaviour*	41	0.00–0.07	0.01 (\pm 0.01)

* Behaviour recorded as percentage of time visible.

indices, the maximum possible value of H also differs. Therefore, we also calculated maximum possible H for each index to allow for relative comparisons between the indices' means and minimum/maximum H values.

Eight different behavioural diversity indices were created to examine the relationship between chimpanzee behavioural diversity and physiological measures (Table 2; see supplementary material to papers published in *Animal Welfare*: <https://www.ufaw.org.uk/the-ufaw-journal/supplementary-material>). Specifically, the first set of indices (1 and 2) only included behaviours that would be considered positive and species-appropriate (many of which are widely accepted, validated indicators of animal welfare, eg play; reviewed in VanderSchuren *et al* 1997; Held & Špinka 2011). Behavioural diversity index 1 splits the behaviours into categories with modifiers (eg Groom-agent, Groom-self, Groom-mutual/multiple), and index 2 combines data with modifiers into the overarching behaviour (eg Groom combined; see Table 2; <https://www.ufaw.org.uk/the-ufaw-journal/supplementary-material>). We split behaviours in order to obtain a fine-grained understanding of how one's role in an interaction (ie agent or recipient) affected welfare outcomes. The second set of indices (3 and 4) included both positive and negative behaviours (eg contact aggression). Index 4 combines the data with modifiers into the overar-

ching behaviour. The third set of indices (5 and 6) included everything in the previous two indices with the addition of human interaction. Index 6 combines the data with modifiers into the overarching behaviour. The final set of indices (7 and 8) removed human interaction and added behaviours that occurred when the focal animal was the recipient. Index 8 combines the data with modifiers into the overarching behaviour. Stereotypic behaviour and inactivity were excluded from all behavioural diversity indices (Swaigood *et al* 2001; Miller *et al* 2011; Razal *et al* 2016). Descriptive statistics and Pearson correlations were used to examine the relationship between all indices.

In order to make comparisons to physiological and behavioural metrics, we calculated mean faecal GCM concentrations, mean IgA concentrations and the mean amount of time spent performing stereotypic behaviour for the three-month period for each individual and used generalised estimating equations (GEE), controlling for institution and sex, to compare these to behavioural diversity. Though the behavioural diversity indices were normally distributed, the faecal GCM, IgA and stereotypic behaviour values were not normally distributed, so GEE was chosen due to the ability to analyse non-normally distributed data without having to make data corrections (Kowalski & Tu 2008; Tang *et al* 2012). The physiological and behavioural measures were the outcome variables in our models. Alpha level was set at $P < 0.05$ for all tests. We also report Cohen's W , which is a measure of effect size (Cohen 1988).

Results

Descriptive statistics for the eight behavioural diversity indices are presented in Table 3. Behavioural diversity ranged from a lowest score of 0.74 for Index 8 and was highest at 2.18 for Index 7. However, the average diversity score only ranged from 1.29 to 1.87. Descriptive statistics for faecal GCM, IgA and stereotypic behaviour are presented in Table 4. Faecal GCM ranged from a lowest concentration of 26.97 ng g⁻¹ and was highest at 1,318.58 ng g⁻¹, with an

Table 5 Generalised estimating equation models comparing the behavioural diversity indices with faecal glucocorticoid metabolites (GCM), immunoglobulin-A (IgA) and stereotypic behaviour.

Analyte	BD index	B	95% CI lower	95% CI upper	Cohen's w	P-value
GCM (ng g ⁻¹)	1	-461.90	-903.02	-20.79	0.32	0.040*
	2	-496.63	-1,038.13	44.87	0.28	0.072
	3	-478.01	-934.88	-21.133	0.32	0.040*
	4	-579.82	-1,106.21	-53.43	0.34	0.031*
	5	-382.89	-748.61	-17.16	0.32	0.040*
	6	-339.296	-773.93	95.33	0.24	0.126
	7	-331.42	-680.80	17.956	0.29	0.063
	8	-715.67	-1,324.58	-106.76	0.36	0.021*
IgA (µg g ⁻¹)	1	87.78	-68.20	243.77	0.17	0.270
	2	732.01	-205.48	1,669.49	0.24	0.126
	3	10.54	-84.02	105.10	0.03	0.827
	4	535.09	-215.81	1,285.99	0.22	0.163
	5	149.09	-77.54	375.72	0.20	0.197
	6	608.73	-168.31	1,385.77	0.24	0.125
	7	-80.996	-227.50	65.51	0.17	0.279
	8	562.64	-192.04	1,317.33	0.23	0.144
Stereotypic behaviour [†]	1	-0.026	-0.050	-0.002	0.34	0.032*
	2	-0.029	-0.056	-0.002	0.33	0.034*
	3	-0.021	-0.044	0.003	0.27	0.082
	4	-0.022	-0.050	0.006	0.24	0.128
	5	-0.019	-0.041	0.003	0.27	0.087
	6	-0.019	-0.043	0.005	0.24	0.119
	7	-0.016	-0.038	0.006	0.23	0.146
	8	-0.017	-0.041	0.007	0.21	0.169

All *P*-values from generalised estimating equation models corrected for sex and institution. For all tests, *df* = 1, and significant differences are indicated at * *P* < 0.05. B: β coefficient; CI is Confidence Interval. † Behaviour recorded as percentage of time visible.

average value of 188.45 ng g⁻¹. IgA ranged from a lowest concentration of 10.29 µg g⁻¹ and was highest at 1,963.89 µg g⁻¹, with an average value of 72.67 µg g⁻¹. Stereotypic behaviour ranged from a lowest value of 0.00 percent time visible to a high of 0.07 percent time visible, with an average of 0.01 percent time visible. Overall, there was a significant correlation between all of the indices, regardless of which behaviours were included (*P* < 0.01 for all comparisons). When the indices comprised of split, fine-grained behaviours (Indices 1, 3, 5 and 7) were compared to one another, the correlation coefficients ranged from 0.931 to 0.980. When the indices comprised of combined behaviours (Indices 2, 4, 6, 8) were compared to one another, the correlation coefficients ranged from 0.923 to 0.980. The correlation coefficients were lower, ranging from 0.445 to 0.606, when comparing indices comprised of split, fine-

grained behaviours to indices comprised of combined behaviours. Significant inverse relationships were found between faecal GCM and five of the eight indices (Table 5). Additionally, we found an inverse relationship between stereotypic behaviour and two indices (Indices 1 and 2). However, there were no significant relationships between any of the indices and IgA concentrations.

Discussion

Our findings provide preliminary evidence that behavioural diversity, and specifically the Shannon Diversity Index (H), can be a valid indicator of animal welfare for professionally managed chimpanzees. All indices were significantly correlated to one another, though it should be noted that there was considerable overlap in terms of the behaviours included in these models. Our analyses indicated that a

majority of the behavioural diversity indices (5 of 8) had significant negative relationships with faecal GCM concentrations. Therefore, subjects who expressed a variety of behaviours from their species-typical repertoire had more regulated adrenal systems, and were likely experiencing less distress and relatively better welfare, than those with lower behavioural diversity. As our data are averaged across a three-month time-period, we cannot speak to differences in acute or chronic stress, merely that overall elevated levels of faecal GCM — which are reliably associated with negative outcomes — are inversely related to behavioural diversity. This is consistent with research linking higher glucocorticoid levels with various environmental or social variables (eg presence of zoo visitors, husbandry) that resulted in health issues and behavioural problems such as hiding, fur-plucking, and pacing (Terio & Munson 2000; Kalthoff *et al* 2001; Schatz & Palme 2001; Wielebnowski *et al* 2002; Lane 2006; Mormède *et al* 2007). This negative relationship between faecal GCM and behavioural diversity is also consistent with studies that have shown that providing additional behavioural opportunities (eg offering environmental enrichment) reduced glucocorticoid levels and behavioural problems (Carlstead *et al* 1993; Boinski *et al* 1999). Our findings regarding faecal GCM and behavioural diversity provide support for using the Shannon Diversity Index as a valid measure of positive welfare in professionally managed chimpanzees.

We did not find significant relationships between behavioural diversity and our other physiological measure, IgA concentrations. To date, the use of IgA in behavioural studies is extremely limited, so this physiological measure warrants additional research as to how it may relate to animal welfare. The potential for the immune system to be activated by a variety of factors, and to reflect both positive and negative events — as well as acute and chronic experiences — may explain why our study did not show a relationship between behavioural diversity and concentrations of IgA.

We also examined the relationship between behavioural diversity and stereotypic behaviour, a commonly used indicator of poor welfare in professionally managed chimpanzees. Stereotypic behaviour (ie repetitive movements such as eye-poking, head-bobbing, pacing and rocking) was inversely related to two indices — Index 1 and Index 2. These happen to be the more traditional indices, which were comprised of a wide range of positive species-typical behaviours. Our results are consistent with those of previous studies which have identified significant inverse relationships between behavioural diversity and stereotypic behaviours (eg Stolba *et al* 1983; Gunn & Morton 1995; reviewed by Miller *et al* 2020).

As this was an exploratory study, we were interested in specifically which combination of variables in the indices would relate to the physiological and behavioural measures for zoo-housed chimpanzees. In a managed setting, some animals have the opportunity to interact frequently with caregivers and/or guests. Historically, measures of behavioural diversity have not incorporated this category (Swaisgood *et al* 2001;

Miller *et al* 2011; Razal *et al* 2016). We chose to include it given that, for certain species, having positive interactions with human caregivers can have positive effects on behaviour (laboratory primates: Baker 1997; Bloomsmith *et al* 1997; Bayne 2002; Waitt *et al* 2002; African and Asian elephants [*Loxodonta africana*] and [*Elephas maximus*]: Greco *et al* 2016; Carlstead *et al* 2019), physiology (clouded leopard [*Neofelis nebulosa*]: Wielebnowski *et al* 2002; black and white rhinoceros [*Diceros bicornis*] and [*Ceratotherium simum*], respectively: Carlstead & Brown 2005; polar bears [*Ursus maritimus*]: Shepherdson *et al* 2013), and even reproductive success (*Felis* spp; Mellen 1991). In our comparisons with faecal GCMs, the addition of human interaction to the behavioural diversity index resulted in a non-significant model when the behaviours were combined (Index 6), but a significant model when the behaviours were more fine-scale (Index 5). Additional research should examine this discrepancy, as one might expect — based on the human-animal relationship literature — that those interactions could be beneficial for individuals of particular species (Hosey 2008; Claxton 2011; Hosey & Melfi 2012, 2014; Melfi 2013). Neither of the indices that included human interaction were associated with IgA or stereotypic behaviour.

Although traditional models of behavioural diversity do not incorporate recipient behaviour, we considered the possibility that being the recipient of an action may be relevant, as the welfare of that individual may be impacted. For example, an individual is calculated as having the same behavioural diversity whether they spend a certain proportion of their day receiving grooming or receiving aggression, yet the welfare implications are considerably different: primates, especially subordinates who are disproportionately victims of aggression and receive less social support (eg grooming), have higher basal cortisol levels relative to higher-ranked group-mates (Abbott *et al* 2003). Indeed, for chimpanzees, grooming helps to solidify and maintain social bonds, and is considered the ‘glue’ that holds the troop together (Langergraber *et al* 2009; Lehmann & Boesch 2009; Mitani 2009; Crockford *et al* 2013). In the current study, the addition of recipient behaviour was only associated with faecal GCM concentrations when behaviours were combined (Index 8) and not when the behaviours were more fine-scale (Index 7). Future studies should examine this discrepancy. Neither of the indices that included recipient behaviour were associated with IgA or stereotypic behaviour.

With the majority of the indices being significantly related to faecal GCM concentrations, two indices being inversely associated with stereotypic behaviour, and high correlations existing between the indices, there is evidence that behavioural diversity — and specifically the H index — can be a robust indicator of positive welfare for professionally managed chimpanzees. We acknowledge that behavioural diversity may be impacted by numerous factors. For instance, individuals may not have an equal opportunity to engage in certain behaviours due to a variety of factors including enclosure size, enrichment schedules, and group

size. Indeed, an individual who is housed in a large enclosure, lives with many conspecifics of various age/sex classes and has regular access to novel enrichment would be expected to have a different behavioural profile than an individual housed alone in a smaller, less-stimulating environment. In fact, the same individual would be expected to spend its time differently when housed in these two environments. However, in the current study, behavioural diversity had a significant negative relationship with faecal GCM across varying environments which suggests it is a potentially robust indicator of positive welfare for chimpanzees.

Animal welfare implications

Overall, this exploratory study of zoo-housed chimpanzees demonstrates some of the complexity around using behavioural diversity as an indicator of positive welfare. Given that some models were not significant, it is necessary for welfare scientists to examine which behavioural measures should (or should not) be included in diversity indices for a particular species, as well as whether behaviours should be split into modifiers or combined for analysis. Moving forward, we expect that animal care professionals hoping to apply behavioural diversity to a wide range of species will find the more traditional models — and especially Index 1 — to be the most useful. Index 1, which was comprised of species-typical behaviours performed by the focal animal (and did not integrate recipient behaviours, negative behaviours, or behaviours that involved human interaction), was the only index that was significantly correlated with both faecal GCM and stereotypic behaviour. Index 1 also did not combine related behaviours (eg ‘Play-Social’ and ‘Play-Solitary’ were entered into the model separately). Moving forward, we must continue to examine behavioural diversity — both for chimpanzees and other taxa — to better understand its potential as a measure of welfare. Future publications should explicitly identify and justify their process for including or excluding behaviours from their calculations to aid in interpreting results. As with any indicator of animal welfare, it is crucial to ensure its validity while incorporating it into welfare assessments and to integrate other validated indicators such as stereotypy (Mason & Latham 2004; Mason 2006; Fureix *et al* 2016) and some forms of inactivity (Fureix & Meagher 2015). We hope that our exploratory process will be helpful in informing these discussions.

Acknowledgements

Thank you to Steve Ross and Mollie Bloomsmith for their endorsement of this study and support throughout. Special thanks to all of the staff and volunteers at participating institutions for their data collection efforts: Busch Gardens Tampa Bay, Chattanooga Zoo, Detroit Zoo, Houston Zoo, John Ball Zoo, Lincoln Park Zoo, Lion Country Safari, Los Angeles Zoo, The Maryland Zoo in Baltimore, North Carolina Zoo, Oklahoma City Zoo, Oregon Zoo, Potawatomi Zoo, Sedgwick County Zoo, St Louis Zoo, Zoo

Knoxville. CZS thanks all of the dedicated volunteers, docents, King Conservation Science Scholars, and Project Search externs that helped code videos, weigh samples in the lab, and move the project forward: Adam, Alia, Alison, Amanda I, Amanda K, Amy, Angel, Betzy, Bob, Bryaene, Carlos, Charles, Chris, Emily, Genny, Gloria, Grace, Harleena, Isabella, Jackie, Jean, Jeff, John, Karen, Lauren, Madeleine, Marcia, Margaret, Mariah, Matt, Meagan, Mike, Nate, Nicole, Paula, Rachel, Rose, Ruben, Sashi, Tanya, Victoria and Violet.

References

- Abbott DH, Keverne EB, Bercovitch FB, Shively CA, Mendoza SP, Saltzman W, Snowdon CT, Ziegler TE, Banjevic M, Garland Jr T and Sapolsky RM** 2003 Are subordinates always stressed? A comparative analysis of rank differences in cortisol levels among primates. *Hormones and Behavior* 43(1): 67-82. [https://doi.org/10.1016/s0018-506x\(02\)00037-5](https://doi.org/10.1016/s0018-506x(02)00037-5)
- Baker KC** 1997 Human interaction as enrichment for captive chimpanzees: A preliminary report. *American Journal of Primatology* 42(2): 92
- Bashaw MJ, Sicks F, Palme R, Schwarzenberger F, Tordiffe AS and Ganswindt A** 2016 Non-invasive assessment of adrenocortical activity as a measure of stress in giraffe (*Giraffa camelopardalis*). *BMC Veterinary Research* 12(1): 235. <https://doi.org/10.1186/s12917-016-0864-8>
- Bayne K** 2002 Development of the human-research animal bond and its impact on animal well-being. *ILAR Journal* 43(1): 4-9. <https://doi.org/10.1093/ilar.43.1.4>
- Bishop NC and Gleeson M** 2009 Acute and chronic effects of exercise on markers of mucosal immunity. *Frontiers in Bioscience* 14: 4444-4456. <https://doi.org/10.2741/3540>
- Blecha F** 2000 *The Biology of Animal Stress* pp 111-122. CABI: Wallingford, UK
- Bloomsmith MA, Lambeth SP, Stone AM and Laule GE** 1997 Comparing two types of human interaction as enrichment for chimpanzees. *American Journal of Primatology* 42(2): 96
- Boinski S, Swing SP, Gross TS and Davis JK** 1999 Environmental enrichment of brown capuchins (*Cebus apella*): Behavioral and plasma and fecal cortisol measures of effectiveness. *American Journal of Primatology* 48(1): 49-68. [https://doi.org/10.1002/\(sici\)1098-2345\(1999\)48:1<49::aid-ajp4>3.0.co;2-6](https://doi.org/10.1002/(sici)1098-2345(1999)48:1<49::aid-ajp4>3.0.co;2-6)
- Butterworth A, Mench J and Wielebnowski N** 2011 Practical strategies to assess (and improve) welfare. In: Appleby MC, Mench JA and Olsson IAS and Hughes BO (eds) *Animal Welfare* pp 200-214. CABI: Wallingford, UK. <https://doi.org/10.1079/9781845936594.0200>
- Carlstead K and Brown JL** 2005 Relationships between patterns of fecal corticoid excretion and behavior, reproduction, and environmental factors in captive black (*Diceros bicornis*) and white (*Ceratotherium simum*) rhinoceros. *Zoo Biology* 24(3): 215-232. <https://doi.org/10.1002/zoo.20050>
- Carlstead K, Brown JL and Seidensticker J** 1993 Behavioral and adrenocortical responses to environmental changes in leopard cats (*Felis bengalensis*). *Zoo Biology* 12(4): 321-331. <https://doi.org/10.1002/zoo.1430120403>

- Carlstead K, Paris S and Brown JL** 2019 Good keeper-elephant relationships in North American zoos are mutually beneficial to welfare. *Applied Animal Behaviour Science* 211: 103-111. <https://doi.org/10.1016/j.applanim.2018.11.003>
- Chinnadurai SK, Millspaugh JJ, Matthews WS, Canter K, Slotow R, Washburn BE and Woods RJ** 2009 Validation of fecal glucocorticoid metabolite assays for South African herbivores. *The Journal of Wildlife Management* 73(6): 1014-1020. <https://doi.org/10.2193/2008-430>
- Claxton AM** 2011 The potential of the human-animal relationship as an environmental enrichment for the welfare of zoo-housed animals. *Applied Animal Behaviour Science* 133(1-2): 1-10. <https://doi.org/10.1016/j.applanim.2011.03.002>
- Cohen J** 1988 *Statistical Power Analysis for the Behavioral Sciences, Second Edition*. Erlbaum: Hillsdale, USA
- Crockford C, Wittig RM, Langergraber K, Ziegler TE, Zuberbühler K and Deschner T** 2013 Urinary oxytocin and social bonding in related and unrelated wild chimpanzees. *Proceedings of the Royal Society B: Biological Sciences* 280: 20122765. <https://doi.org/10.1098/rspb.2012.2765>
- Cronin KA and Ross SR** 2019 Technical contribution: A cautionary note on the use of behavioural diversity (H-Index) in animal welfare science. *Animal Welfare* 28: 157-164. <https://doi.org/10.7120/09627286.28.2.157>
- Duncan IJ and Petherick JC** 1991 The implications of cognitive processes for animal welfare. *Journal of Animal Science* 69: 5017-5022. <https://doi.org/10.2527/1991.69125017x>
- Elsasser TH, Klasing KC, Filipov N and Thompson F** 2000 *The Biology of Animal Stress* pp 77-110. CAB: Wallingford, UK
- Eriksson E, Royo F, Lyberg K, Carlsson HE and Hau J** 2004 Effect of metabolic cage housing on immunoglobulin A and corticosterone excretion in faeces and urine of young male rats. *Experimental Physiology* 89: 427-433. <https://doi.org/10.1113/expphysiol.2004.027656>
- Fraser D and Duncan IJH** 1998 'Pleasures', 'pains' and animal welfare: Toward a natural history of affect. *Animal Welfare* 7: 383-396
- Friard O and Gamba M** 2016 BORIS: a free, versatile open-source event-logging software for video/audio coding and live observations. *Methods in Ecology and Evolution* 7: 1325-1330. doi:10.1111/2041-210X.12584
- Fureix C and Meagher RK** 2015 What can inactivity (in its various forms) reveal about affective states in non-human animals? A review. *Applied Animal Behaviour Science* 171: 8-24. <https://doi.org/10.1016/j.applanim.2015.08.036>
- Fureix C, Walker M, Harper L, Reynolds K, Saldivia-Woo A and Mason G** 2016 Stereotypic behaviour in standard non-enriched cages is an alternative to depression-like responses in C57BL/6 mice. *Behavioural Brain Research* 305: 186-190. <https://doi.org/10.1016/j.bbr.2016.02.005>
- Gourkow N, Hamon SC and Phillips CJ** 2014 Effect of gentle stroking and vocalization on behavior, mucosal immunity and upper respiratory disease in anxious shelter cats. *Preventative Veterinary Medicine* 117: 266-275. <https://doi.org/10.1016/j.prevetmed.2014.06.005>
- Gourkow N and Phillips CJ** 2015 Effect of interactions with humans on behavior, mucosal immunity and upper respiratory disease of shelter cats rated as contented on arrival. *Preventative Veterinary Medicine* 121: 288-296. <https://doi.org/10.1016/j.prevetmed.2015.07.013>
- Gourkow N and Phillips CJ** 2016 Effect of cognitive enrichment on behavior, mucosal immunity and upper respiratory disease of shelter cats rated as frustrated on arrival. *Preventative Veterinary Medicine* 131: 103-110. <https://doi.org/10.1016/j.prevetmed.2016.07.012>
- Greco BJ, Meehan CL, Hogan JN, Leighty KA, Mellen J, Mason GJ and Mench JA** 2016 The days and nights of zoo elephants: using epidemiology to better understand stereotypic behavior of African elephants (*Loxodonta africana*) and Asian elephants (*Elephas maximus*) in North American zoos. *PLoS One* 11: e0144276. <https://doi.org/10.1371/journal.pone.0144276>
- Guhad FA and Hau J** 1996 Salivary IgA as a marker of social stress in rats. *Neuroscience Letters* 216: 137-140. [https://doi.org/10.1016/0304-3940\(96\)13037-8](https://doi.org/10.1016/0304-3940(96)13037-8)
- Gunn D and Morton DB** 1995 Inventory of the behaviour of New Zealand White rabbits in laboratory cages. *Applied Animal Behaviour Science* 45: 277-292. [https://doi.org/10.1016/0168-1591\(95\)00627-5](https://doi.org/10.1016/0168-1591(95)00627-5)
- Heintz MR, Santymire RM, Parr LA and Lonsdorf EV** 2011 Validation of a cortisol enzyme immunoassay and characterization of salivary cortisol circadian rhythm in chimpanzees (*Pan troglodytes*). *American Journal of Primatology* 73: 903-908. <https://doi.org/10.1002/ajp.20960>
- Held SD and Špinka M** 2011 Animal play and animal welfare. *Animal Behaviour* 81: 891-899. <https://doi.org/10.1016/j.anbehav.2011.01.007>
- Hosey G** 2008 A preliminary model of human-animal relationships in the zoo. *Applied Animal Behaviour Science* 109: 105-127. <https://doi.org/10.1016/j.applanim.2007.04.013>
- Hosey G and Melfi V** 2012 Human-animal bonds between zoo professionals and the animals in their care. *Zoo Biology* 31: 13-26. <https://doi.org/10.1002/zoo.20359>
- Hosey G and Melfi V** 2014 Human-animal interactions, relationships and bonds: a review and analysis of the literature. *International Journal of Comparative Psychology* 27: 117-142. <https://doi.org/10.46867/ijcp.2014.27.01.01>
- Kalthoff A, Schmidt C and Sachser N** 2001 Influence of zoo visitors on behaviour and salivary corticosterone concentrations of zoo animals. *KTBL SCHRIFT* 403: 104-112
- Kowalski J and Tu XM** 2008 *Modern Applied U-statistics, Volume 714*. John Wiley and Sons: Hoboken, NJ, USA
- Lane J** 2006 Can non-invasive glucocorticoid measures be used as reliable indicators of stress in animals? *Animal Welfare* 15: 331-342
- Langergraber KE, Mitani JC and Vigilant L** 2009 Kinship and social bonds in female chimpanzees (*Pan troglodytes*). *American Journal of Primatology* 71: 840-851. <https://doi.org/10.1002/ajp.20711>
- Lantz EL, Lonsdorf EV, Heintz MR, Murray CM, Lipende I, Travis DA and Santymire RM** 2016 Non-invasive quantification of immunoglobulin A in chimpanzees (*Pan troglodytes schweinfurthii*) at Gombe National Park, Tanzania. *American Journal of Primatology* 80: e22558. <https://doi.org/10.1002/ajp.22558>
- Lehmann J and Boesch B** 2009 Sociality of the dispersing sex: the nature of social bonds in West African female chimpanzees, *Pan troglodytes*. *Animal Behavior* 77: 377-387. <https://doi.org/10.1016/j.anbehav.2008.09.038>

- Mason G** 2006 Stereotypic behaviour in captive animals: Fundamentals and implications for welfare and beyond. In: Mason G and Rushen J (eds) *Stereotypic Animal Behaviour: Fundamentals and Applications to Welfare* pp 325-367. CAB International: Wallingford, UK. <https://doi.org/10.1079/9780851990040.0325>
- Mason GJ, Cooper J and Clarebrough C** 2001 Frustrations of fur-farmed mink. *Nature* 410: 35. <https://doi.org/10.1038/35065157>
- Mason GJ and Latham NR** 2004 Can't stop, won't stop: is stereotypy a reliable animal welfare indicator? *Animal Welfare* 13: 57-69
- Meagher RK, Campbell DL and Mason GJ** 2017 Boredom-like states in mink and their behavioural correlates: a replicate study. *Applied Animal Behaviour Science* 197: 112-119. <https://doi.org/10.1016/j.applanim.2017.08.001>
- Meagher RK and Mason GJ** 2012 Environmental enrichment reduces signs of boredom in caged mink. *PLoS One* 7(11): e49180. <https://doi.org/10.1371/journal.pone.0049180>
- Melfi V** 2013 Is training zoo animals enriching? *Applied Animal Behaviour Science* 147: 299-305. <https://doi.org/10.1016/j.applanim.2013.04.011>
- Mellen JD** 1991 Factors influencing reproductive success in small captive exotic felids (*Felis* spp): a multiple regression analysis. *Zoo Biology* 10: 95-110. <https://doi.org/10.1002/zoo.1430100202>
- Miller LJ, Mellen J, Greer T and Kuczaj SA** 2011 The effects of education programmes on Atlantic bottlenose dolphin (*Tursiops truncatus*) behaviour. *Animal Welfare* 20: 159-172
- Miller LJ, Pisacane CB and Vicino GA** 2016 Relationship between behavioural diversity and faecal glucocorticoid metabolites: a case study with cheetahs (*Acinonyx jubatus*). *Animal Welfare* 25: 325-329. <https://doi.org/10.7120/09627286.25.3.325>
- Miller LJ, Vicino GA, Sheffel J and Lauderdale LK** 2020 Behavior diversity as a potential indicator of positive animal welfare. *Animals* 10: 1211. <https://doi.org/10.3390/ani10071211>
- Mitani JC** 2009 Male chimpanzees form enduring and equitable social bonds. *Animal Behavior* 77: 633-640. <https://doi.org/10.1016/j.anbehav.2008.11.021>
- Moberg GP** 1987 Influence of the adrenal axis upon the gonads. *Oxford Reviews of Reproductive Biology* 9: 456
- Moberg GP and Mench JA** 2000 *The Biology of Animal Stress: Basic Principles and Implications for Animal Welfare*. CAB International: Wallingford, UK. <https://doi.org/10.1079/9780851993591.0000>
- Mormède P, Andanson S, Aupérin B, Beerda B, Guémené D, Malmkvist J, Manteca X, Manteuffel G, Prunet P, van Reenen CG, Richard S and Veissier I** 2007 Exploration of the hypothalamic-pituitary-adrenal function as a tool to evaluate animal welfare. *Physiology and Behavior* 92: 317-339. <https://doi.org/10.1016/j.physbeh.2006.12.003>
- Murray CM, Heintz MR, Lonsdorf EV, Parr LA and Santymire RM** 2013 Validation of a field technique and characterization of fecal glucocorticoid metabolite analysis in wild chimpanzees (*Pan troglodytes*). *American Journal of Primatology* 75: 57-64. <https://doi.org/10.1002/ajp.22078>
- Palme R, Touma C, Arias N, Dominchin MF and Lepschy M** 2013 Steroid extraction: getting the best out of faecal samples. *Wien Tierarztl Monatsschr* 100: 238-246
- Pressman SD and Cohen S** 2005 Does positive affect influence health? *Psychological Bulletin* 131: 925. <https://doi.org/10.1037/0033-2909.131.6.925>
- Razal CB, Pisacane CB and Miller LJ** 2016 Multifaceted approach to personality assessment in cheetahs (*Acinonyx jubatus*). *Animal Behavior and Cognition* 3: 22-31. <https://doi.org/10.12966/abc.02.02.2016>
- Renner MJ and Lussier JP** 2002 Environmental enrichment for the captive spectacled bear (*Tremarctos ornatus*). *Pharmacology Biochemistry and Behavior* 73: 279-283. [https://doi.org/10.1016/s0091-3057\(02\)00786-4](https://doi.org/10.1016/s0091-3057(02)00786-4)
- Rivier C** 1995 Luteinizing-hormone-releasing hormone, gonadotropins, and gonadal steroids in stress. *Annals of the New York Academy of Sciences* 771: 187-191. <https://doi.org/10.1111/j.1749-6632.1995.tb44680.x>
- Ross SR and Lukas KE** 2001 The care and management of captive chimpanzees workshop: Managing social behavior. *Journal of Applied Animal Welfare Science* 4: 299-301. https://doi.org/10.1207/s15327604jaws0404_07
- Rushen J, Lawrence AB and Terlouw EMC** 1993 The motivational basis of stereotypies. In: Lawrence A and Rushen J (eds) *Stereotypic Animal Behaviour* pp 41-64. CAB International: Wallingford, UK
- Sapolsky RM** 2002 Endocrinology of the stress-response. In: Becker J, Breedlove S, Crews D and McCarthy M (eds) *Behavioral Endocrinology, Second Edition* pp 409-450. MIT Press: Cambridge, MA, USA
- Sarrafcchi A and Blokhuis HJ** 2013 Equine stereotypic behaviors: causation, occurrence, and prevention. *Journal of Veterinary Behavior: Clinical Applications and Research* 8: 386-394. <https://doi.org/10.1016/j.jveb.2013.04.068>
- Schatz S and Palme R** 2001 Measurement of faecal cortisol metabolites in cats and dogs: a non-invasive method for evaluating adrenocortical function. *Veterinary Research Communications* 25: 271-287
- Shannon CE and Weaver W** 1949 The mathematical theory of communication. *University of Illinois Press* 1, 11: 117
- Shepherdson D, Lewis KD, Carlstead K, Bauman J and Perrin N** 2013 Individual and environmental factors associated with stereotypic behavior and fecal glucocorticoid metabolite levels in zoo housed polar bears. *Applied Animal Behaviour Science* 147: 268-277. <https://doi.org/10.1016/j.applanim.2013.01.001>
- Shepherdson DJ, Carlstead K, Mellen JD and Seidensticker J** 1993 The influence of food presentation on the behavior of small cats in confined environments. *Zoo Biology* 12: 203-216. <https://doi.org/10.1002/zoo.1430120206>
- Siegford JM** 2013 Multidisciplinary approaches and assessment techniques to better understand and enhance zoo nonhuman animal welfare. *Journal of Applied Animal Welfare Science* 16: 300-318. <https://doi.org/10.1080/10888705.2013.827914>
- Simpson EH** 1949 Measurement of diversity. *Nature* 163: 688
- Skandakumar S, Stodulski G and Hau J** 1995 Salivary IgA: a possible stress marker in dogs. *Animal Welfare* 4: 339-350.
- Staley M, Connors MG, Hall K and Miller LJ** 2018 Linking stress and immunity: Immunoglobulin A as a non-invasive physiological biomarker in animal welfare studies. *Hormones and Behavior* 102: 55-68. <https://doi.org/10.1016/j.yhbeh.2018.04.011>

- Stolba A, Wood-Gush DG and Baker N** 1983 The characterisation of stereotyped behaviour in stalled sows by informational redundancy. *Behaviour* 87: 157-182. <https://doi.org/10.1163/156853983x00417>
- Swaigood R and Shepherdson D** 2006 Environmental enrichment as a strategy for mitigating stereotypies in zoo animals: a literature review and meta-analysis. In: Mason G and Rushen J (eds) *Stereotypic Animal Behaviour: Fundamentals and Applications to Welfare, Second Edition* pp 256-285. CABI: Wallingford, UK. <https://doi.org/10.1079/9780851990040.0256>
- Swaigood RR, White AM, Zhou X, Zhang G and Lindburg DG** 2005 How do giant pandas (*Ailuropoda melanoleuca*) respond to varying properties of enrichments? A comparison of behavioral profiles among five enrichment items. *Journal of Comparative Psychology* 119: 325. <https://doi.org/10.1037/0735-7036.119.3.325>
- Swaigood RR, White AM, Zhou X, Zhang H, Zhang G, Wei R, Hare VJ, Tepper EM and Lindburg DG** 2001 A quantitative assessment of the efficacy of an environmental enrichment programme for giant pandas. *Animal Behaviour* 61: 447-457. <https://doi.org/10.1006/anbe.2000.1610>
- Tang W, He H and Tu X** 2012 *Applied Categorical and Count Data Analysis*. Chapman and Hall/CRC: Boca Raton, FL, USA
- Terio KS and Munson L** 2000 Gastritis in cheetahs and relatedness to adrenal function. Felid Taxon Advisory Group Action Plan. American Zoo and Aquarium Association: Wheeling, Virginia, USA
- Tsujita S and Morimoto K** 1999 Secretory IgA in saliva can be a useful stress marker. *Environmental Health and Preventive Medicine* 4: 1-8. <https://doi.org/10.1007/bf02931243>
- VanderSchuren LJM, Niesink RJM and van Ree JM** 1997 The neurobiology of social play behavior in rats. *Neuroscience & Biobehavioral Reviews* 21: 309-326. [https://doi.org/10.1016/s0149-7634\(96\)00020-6](https://doi.org/10.1016/s0149-7634(96)00020-6)
- Vickery S and Mason G** 2004 Stereotypic behavior in Asiatic black and Malayan sun bears. *Zoo Biology* 23: 409-430. <https://doi.org/10.1002/zoo.20027>
- Waite C, Buchanan-Smith HM and Morris K** 2002 The effects of caretaker-primate relationships on primates in the laboratory. *Journal of Applied Animal Welfare Science* 5: 309-319. https://doi.org/10.1207/s15327604jaws0504_05
- Watanuki S and Kim YK** 2005 Physiological responses induced by pleasant stimuli. *Journal of Physiological Anthropology and Applied Human Science* 24: 135-138. <https://doi.org/10.2114/jpa.24.135>
- Wemelsfelder F, Haskell M, Mendl MT, Calvert S and Lawrence AB** 2000 Diversity of behaviour during novel object tests is reduced in pigs housed in substrate-impooverished conditions. *Animal Behaviour* 60: 385-394. <https://doi.org/10.1006/anbe.2000.1466>
- Whitham JC and Wielebnowski N** 2013 New directions for zoo animal welfare science. *Applied Animal Behaviour Science* 147: 247-260. <https://doi.org/10.1016/j.applanim.2013.02.004>
- Wielebnowski NC, Fletchall N, Carlstead K, Busso JM and Brown JL** 2002 Non-invasive assessment of adrenal activity associated with husbandry and behavioral factors in the North American clouded leopard population. *Zoo Biology* 21: 77-98. <https://doi.org/10.1002/zoo.10005>
- Würbel H, Stauffacher M and von Holst D** 1996 Stereotypies in laboratory mice: Quantitative and qualitative description of the ontogeny of 'wire-gnawing' and 'jumping' in ICR and Zur: ICR nu. *Ethology* 102: 371-385. <https://doi.org/10.1111/j.1439-0310.1996.tb01133.x>