JZAR JOURNAL OF ZOO AND AQUARIUM RESEARCH
www.jzar.org

Review article

Evidence-based zoo animal welfare assessment: Putting science into practice

Marina Salas[1], Oriol Tallo-Parra[2,3] and Xavier Manteca[2,3]

[1]Antwerp Zoo Centre for Research and Conservation, Royal Zoological Society of Antwerp, Koningin Astridplein 20-26, 2018 Antwerpen, Belgium
[2]School of Veterinary Science, Universitat Autònoma de Barcelona, Campus UAB, 08193 Barcelona, Spain
[3]Animal Welfare Education Centre, AWEC Advisors SL, Universitat Autònoma de Barcelona, Campus UAB, 08193 Barcelona, Spain

Correspondence: Marina Salas, email; marina.salas@kmda.org

Keywords: behaviour, physiology, validation, welfare indicators, zoo animal welfare

Article history:
Received: 16 Nov 2023
Accepted: 08 Jul 2024
Published online: 31 Oct 2024

Abstract

This comprehensive review explores evidence-based strategies for assessing and enhancing animal welfare in modern zoos and aquariums. The two primary objectives are to explore the ways in which understanding behavioural biology and natural history of a given species can enhance zoo animal welfare assessments and discuss how current knowledge of fundamental principles regarding animal behaviour and physiology can help identify and validate welfare indicators. Species-specific protocols, generic protocols and risk assessment methods are examined and the complexities of using natural behaviour as a welfare indicator are explored, acknowledging the inherent challenges of comparing captive and wild behaviours. Behavioural indicators as predominant tools in welfare assessment are analysed for their selection, development and validation. Challenges such as observer bias and external influences are discussed, highlighting the importance of ongoing research and collaboration for refining behavioural indicators. The review extends to physiological indicators, focusing on their diversity and complementarity with behavioural assessments. The selection process involves consideration of species-specific characteristics, biological matrices and sampling methodology. Challenges in the validation of physiological indicators are discussed, underlining the need for comprehensive studies. In conclusion, this review advocates for an integrated, evidence-based approach that combines behavioural and physiological indicators, acknowledging the challenges and offering practical insights for advancing animal welfare in zoo settings.

Introduction

Ensuring the highest possible standards of animal welfare has become an absolute priority for modern zoos and aquaria (hereafter referred to collectively as zoos). First and foremost, animal welfare is grounded in ethical concerns that derive from the fact that animals are sentient beings, i.e. are capable of suffering and experiencing emotions (Le Neindre et al. 2017). Also, providing the best possible conditions for zoo-kept animals is an essential requirement if zoos are to realise their education and conservation functions.

Scientific-based tools to assess animal welfare are needed to identify welfare problems and to monitor progress when improvement strategies are implemented. However, rigorous, science-based assessment of zoo animal welfare is a challenge due both to the sheer diversity of species kept in zoos and the lack of knowledge of the general biology and specific needs of many of them.

The objectives of this review are to discuss how zoo animal welfare assessment can benefit from a sound understanding of both the behavioural biology and natural history of the species of interest and how current knowledge of the fundamental principles of animal behaviour and physiology can assist when identifying and validating behavioural and physiological animal welfare indicators.

Species-specific protocols and risk assessment methods to assess zoo animal welfare

Tallo-Parra et al. (2023) reviewed the main approaches to assessing zoo animal welfare, including species-specific protocols, generic protocols and risk assessment methods, assessment of welfare based on time budgets, keepers' ratings and cognitive bias testing.

Research on animal welfare assessment has made significant progress over recent years, mainly for farm animals. One of the main breakthroughs in farm animal welfare assessment has been the development of species-specific protocols, such as the Welfare Quality© protocols (Botreau et al. 2007), which include four animal welfare principles (feeding, environment, health and behaviour) that coincide with the four physical domains of the Five Domains Model (Mellor 2016). Each principle includes several animal welfare criteria and each criterion is evaluated through one or several indicators. The welfare principles and criteria are the same regardless of the species, whereas the welfare indicators can vary across species.

Although originally developed for farm animals kept under intensive conditions, the Welfare Quality© assessment protocols have been used as a basis to develop protocols for zoo animals, such as bottlenose dolphins *Tursiops truncatus* (Clegg et al. 2015) and dorcas gazelles *Gazella dorcas* (Salas et al. 2018). A slightly different approach was followed by Yon et al. (2019), who developed a behavioural protocol to assess the welfare of captive African and Asian elephants *Loxodonta africana* and *Elephas maximus.*

Species-specific protocols have several advantages over other welfare assessment methods as they are meant to cover all aspects of animal welfare, use measures that have the potential to be tested for validity and reliability and are tailored to the biological needs and peculiarities of each species. However, their main limitation is that very few such protocols have been developed and for most zoo animals there is no specific welfare assessment protocol.

Due in part to the lack of species-specific protocols for most zoo-kept animals, several authors have proposed generic assessment protocols, i.e. protocols that can be used in any species. For example, Sherwen et al. (2018) developed a welfare risk assessment protocol which includes a total of 20 indicators (both animal- and resource-based) as well as a scoring methodology so that each indicator is given a value of 0, 1 or 2. This method is meant to identify potential welfare issues and prioritise improvement actions, so that zoo personnel can take a proactive approach rather than simply flag welfare problems when they have already appeared.

One of the main difficulties of generic assessment protocols, however, is that they require a sound knowledge of the natural history and behaviour of the target species. For example, some of the questions in the assessment protocol developed by Sherwen et al. (2018) refer to the species' adequate social setting or behaviour, meaning that the protocol needs to be tailored to each species' biology, including its natural behaviour. However, using natural behaviour to assess the welfare of animals under human care presents several limitations.

Using natural behaviour to assess zoo animal welfare

The welfare of wild animals under human care is often assessed by comparing their behaviour with that of their free-ranging conspecifics (Veasey et al. 1996). However, such an approach has many methodological and theoretical problems. One—and perhaps the most critical—is that a difference between behaviour in captivity and in the wild does not necessarily imply that animals in captivity have poor welfare, as behaviour is flexible and context-dependent. Moreover, animals in the wild can still have very poor welfare under some circumstances. Veasey et al. (1996) provide a thorough discussion of the limitations of comparing behaviour in the wild and in captivity to assess welfare.

Despite the above considerations, it remains true that the expression of some natural behaviours has obvious welfare benefits and can even be considered a requisite for good welfare. The link between natural behaviour and welfare goes as far back as the Brambell Report, which stated that a degree of confinement that frustrates most of an animal's natural behaviour should be seen as inadequate from an animal welfare standpoint (Brambell 1965). More recently, the relationship between natural behaviour and welfare has been discussed in the context of positive welfare and it has been argued that positive emotions—which are a central feature of positive welfare—are closely linked to the performance of some behaviours such as play, exploration and maternal care (Lawrence et al. 2019).

If it is accepted that the performance of some patterns of natural behaviour is a condition for good welfare, then the question of which behaviours are indeed important follows. Answering this question becomes particularly difficult as there are differences both across species and across environments. Ideally, an experimental approach would be used to find out whether—for each species and each possible environment—the performance of a given behaviour leads to welfare benefits or, alternatively, whether the frustration of such performance has a harmful effect. This approach is not realistic given the sheer diversity of zoo-kept species and the variety of environments in which they are kept. Therefore, a more practical approach should be followed by assuming that those behavioural patterns that have been found to be important for the welfare of several species are likely to be important for the welfare of all zoo-kept animals. Following this idea, exploration, foraging behaviour and social interaction with conspecifics, among others, should be labelled as being relevant for welfare (Bracke and Hopster 2006; Lawrence et al. 2019). It is important to consider that this approach does not mean that a detailed knowledge of the natural behaviour of the target species is not important. On the contrary, each species has its own natural foraging habits and social setting, for example, and knowledge of this is necessary to provide opportunities for the expression of important behaviours.

Identifying and validating behavioural indicators

Behavioural indicators stand out as the most predominant tools in the evaluation of animal welfare (Binding et al. 2020). Their popularity can be attributed to the non-invasive nature of behavioural observations (Hosey et al. 2009), for which advanced technology is typically not required (Binding et al. 2020), and their cost-effective application (Hill and Broom 2009). The most frequently used welfare indicators based on behaviour are those that assess abnormal behaviours and changes in the expression (i.e. frequency, duration or intensity) of normal behaviours (Manteca et al. 2016; Tallo-Parra et al. 2023).

Selecting behavioural indicators

When it comes to assessing animal welfare in a zoo environment, the selection of appropriate behavioural indicators is a critical aspect of the process. These indicators serve as valuable tools for caretakers, researchers and veterinarians. It is important to choose indicators that are relevant, reliable, sensitive and specific to the species and context. Indicators related to abnormal behaviours encompass a range of behaviours, including abnormal repetitive behaviours, damaging behaviours (e.g. self-injurious behaviours, regurgitation and reingestion) and apathy (Manteca et al. 2016; Tallo-Parra et al. 2023).

Evaluating changes in the expression of normal behaviours can provide valuable insights into animal welfare (Manteca et al.

2016; Tallo-Parra et al. 2023). These behaviours include affiliative, agonistic (or aggressive), maternal and play behaviours, as well as activities such as food intake, rumination, sleep behaviour, anticipatory behaviour, use of the enclosure, displacement behaviours and more. Alterations in the expression of these behaviours can potentially indicate welfare concerns.

A comprehensive understanding of the species being assessed, accounting for individual variations, is essential to identify a normal range of activity and deviations from this. Unfortunately, for several species housed in zoos, such knowledge is lacking as their needs in natural habitats are not always fully understood (Hill and Broom 2009). This gap in understanding highlights the importance of ongoing research and collaboration among experts to refine and expand selection of behavioural indicators, ultimately enhancing welfare assessment methods in zoo settings.

Before initiating behavioural observations, it is important to consider various factors, as outlined in Table 1.

Developing an ethogram

To study animal behaviour, an ethogram must be created. This is a comprehensive list of behaviours exhibited by a species or an individual. A complete ethogram provides a thorough insight into the behavioural repertoire of a species. The extent and categorisation of behaviours within the ethogram depend on the research or behavioural question at hand (e.g. What is the activity budget of the animal? What feeding-related behaviours does this animal display?). The behaviours can be grouped into categories such as social behaviours, locomotion, feeding, resting and more. Additionally, behaviours can be classified as either state (long in duration, like resting) or event (short in duration, like yawning).

Each behaviour in the ethogram should have a clear and objective definition, avoiding ambiguity or subjectivity. This ensures consistent behaviour identification and recording by observers without subjective interpretation. Behaviours should be mutually exclusive. The ethogram should also be complete and exhaustive (i.e. by adding categories such as 'other: any behaviour not described by other definitions' or 'out of view: the animal is not visible to the observer').

Ethograms should be dynamic documents, open to revision as more is learned about the species. Researchers should regularly update and refine ethograms based on new observations and insights. Different species have distinct natural behaviours and specific behavioural needs, requiring customised, species-specific behavioural indicators (Browning 2023). Given the lack of knowledge surrounding behaviours in several species of zoo-housed animals, collaborative efforts between researchers and experts in animal behaviour are crucial for developing species-specific ethograms.

Behavioural observations

Animal observations can be conducted at either the individual or group level. The three most common methods of behavioural observations used in zoos are all-occurrence, intervals and continuous (Altmann 1974). Depending on the research question, one or a combination of methods is chosen. For brief behaviours like certain social interactions or vocalisations, all-occurrence might be preferred, where all selected events observed within a specified time are recorded, providing the frequency of that particular behaviour per observed time. For lengthier behaviours such as foraging or ruminating, intervals (recording behaviour at pre-scheduled points in time, providing a percentage of time) or continuous (recording all behaviours displayed during an observation period, with metrics including duration, frequency or percentage of time) may be more suitable.

However, it is not as straightforward as it might seem, as the importance of a behaviour is not solely determined by its frequency and/or duration; the intensity of the behaviour is also crucial in certain situations. For instance, an aggressive interaction may exhibit various levels of intensity (Manteca et al. 2016).

Conducting longitudinal studies, whenever possible, is very valuable. These studies enable the tracking of changes in behaviour over time, offering insights into how alterations (e.g. in the individuals' environment, social group or health) impact their behaviour and consequently their welfare. This approach is essential for monitoring the effects of improvements or changes in husbandry, enclosures or other aspects related to the animal's life.

Table 1. Examples of pre-observation considerations for behavioural studies

Element	Examples of considerations
Research or behavioural question	What is the learning objective of these observations? Will the outcome be used to make evidence-based decisions? What is the behavioural question?
Ethogram	What behaviours should be included in the ethogram to answer the research question?
Individual or group	Which animal or groups of animals need to be observed to answer the research question?
Methodology	When is the optimal timing for observations to effectively address the research question? What methodology be employed for recording and documenting observed behaviours?

When using normal behaviours as indicators of welfare, it is essential to closely monitor changes in their expression, specifically alterations in the frequency, intensity and/or duration of these behaviours (Manteca et al. 2016; Tallo-Parra et al. 2023). Consequently, regular monitoring of normal animal behaviour becomes imperative to identify and address changes in the expression of these behaviours.

Validating behavioural indicators

After collecting behavioural data, the next step involves delving into the analysis process, essential for addressing the research or behavioural questions. The gathered data can be compared to a predefined target, providing a reference for assessing deviations or patterns. Exploring individual variations within the data allows for nuanced observations and the identification of unique behavioural profiles among individuals. Analysing the temporal aspects of behaviour, especially considering the durations or frequencies of specific behaviours, can be useful for detecting patterns or changes over time or after an improvement has been made.

The welfare of zoo-housed wild animals is sometimes evaluated by comparing their behaviour to that of their conspecifics in the wild (Veasey et al. 1996). However, as discussed earlier, this approach presents methodological and theoretical challenges.

Behavioural indicators can be validated by, for instance, comparing them to established welfare standards or measures, which refer to established criteria or guidelines used to assess the welfare of animals. However, as mentioned earlier, there is a lack of knowledge concerning zoo animals, often resulting in the absence of welfare standards or measures for comparison. Therefore, other methods are used for validating behavioural indicators, including the following. 1) Expert consensus uses insights of experts familiar with the species' natural history and behaviour to detect which behaviours can be used to assess animal welfare. This consensus can help establish a preliminary set of indicators to be further tested. 2) Inter-observer reliability tests ensure that different observers can consistently identify and record the same behaviours, which is crucial for reliability. This involves training multiple observers to independently record behaviours. High inter-observer reliability scores indicate that the behaviours are well-defined and consistently recognised across different observers. 3) Correlation with physiological measures can be explored by validating behavioural indicators and comparing them with physiological measures of welfare, such as cortisol levels, heart rate or immune function. This correlation helps confirm that the observed behaviours are accurate reflections of the animals' physiological state. 4) Experimental validation involves conducting controlled experiments where specific variables are manipulated to help validate behavioural indicators. By observing changes in behaviour in response to these manipulations, it can be determined whether the indicators accurately reflect welfare changes. For instance, introducing a new form of environmental enrichment might increase exploratory behaviours, indicating improved welfare. 5) Longitudinal studies track individual animals over extended periods, allowing assessment of the stability and reliability of behavioural indicators over time. Consistent behavioural patterns across different life stages and varying conditions strengthen the validity of these indicators.

Challenges and limitations

Although behavioural indicators are widely recognised as the most popular indicators for evaluating animal welfare (Binding et al. 2020), they come with inherent challenges and limitations. Observer bias is a significant issue, where human observers may unintentionally introduce bias through subjective interpretation and expectation. Variations in observer training, experience and fatigue can lead to inconsistent data, which can be mitigated through rigorous training, periodic reliability checks and the use of multiple observers to cross-verify data. Temporal variability is another challenge, as animal behaviour can vary significantly throughout the day and across different seasons, influenced by factors such as weather condition. Longitudinal studies and continuous monitoring can help account for these variations. Environmental influences also impact behaviour; the design and complexity of enclosures, presence of environmental enrichment items and overall environment can cause differences in behaviour. Additionally, the presence of caretakers and their interactions with animals can alter behaviour through direct engagement or anticipation of activities, necessitating observations that minimise caretaker influence for unbiased data. Advances in technology such as remote behavioural monitoring and/or automated tracking systems can help mitigate some of these limitations (Diana et al. 2021). However, there are also technological limitations, including technical issues, calibration errors and high initial setup costs. Moreover, not all species or behaviours are easily tracked with current technology, presenting further challenges. Different zoos and animal care institutions may have varying standards, practices and cultural attitudes towards animal welfare, influencing the implementation and validation of behavioural indicators and making it challenging to develop universally applicable standards.

Identifying and validating physiological indicators

In welfare science, the term 'physiological indicators' is typically employed to encompass and represent a broad and highly diverse group of physiological, neuroendocrine, haematological, cardiovascular, respiratory, immunological and/or cellular welfare-related biomarkers or changes among others (Broom and Johnson 2019). The physiological indicators related to the stress response are particularly interesting for zoo-housed animals and are the most employed physiological indicators (Tarlow and Blumstein 2007). However, it is important to acknowledge that the relationship between each physiological indicator and the stress response is different (Dickens and Romero 2013) and that many other physiological indicators not directly related to stress can also be very interesting options (Tallo-Parra et al. 2023).

Most physiological indicators must be extracted from a biological matrix and quantified in a laboratory after several laboratory procedures. This section focuses on this type of physiological indicator despite most of the information also being valid for other physiological indicators.

The high diversity and complementarity of physiological indicators makes them an interesting option to use in combination with behavioural and other indicators when assessing welfare (Staley et al. 2018; Whitham and Wielebnowski 2013). However, their diversity and complexity also imply that meeting the necessary requirements to correctly use and interpret physiological indicators is of upmost importance and usually challenging. The use of physiological indicators requires in-depth knowledge of the characteristics of each indicator and biological matrix, individual, species and context, in conjunction with an appropriate sampling design and execution. Missing important information about some of the previous features may lead to very difficult or incorrect interpretations of physiological indicators, thereby impacting the validity of the welfare evaluation (Ralph and Tilbrook 2016). Three elements are critical to correctly integrate the use of physiological indicators in the assessment of welfare in zoo-housed animals: knowing how to identify the right physiological indicators for each welfare assessment, acknowledging the importance of adequate validations for physiological indicators and understanding the challenges associated with their use in non-validated species, matrices and/or contexts.

Many considerations should be taken into account when selecting a physiological indicator for a welfare assessment. Examples of these are shown in Table 2. In addition to the elemental considerations related to the indicator (e.g. which welfare aspect it represents and how), other very important aspects should be contemplated such as the characteristics of the indicator's biological matrix of origin, as well as the sampling possibilities that the target species and particular individuals and contexts allow (Gormally and Romero 2020).

Importantly, many 'secondary elements' that are not the indicator itself, such as the matrix or the sampling design, exert a profound influence on the indicator's capacity to represent a welfare-related physiological status. They do so to such an extent as to modify the indicator's sensitivity to welfare changes or to define the timeframe it can provide information about (Sadoul and Geffroy 2019; Tallo-Parra et al. 2023). For instance, glucocorticoid concentrations assessed in different matrices (such as blood, faeces or hair) can provide different information about the physiological state and well-being of the same individual (Gormally and Romero 2020). Therefore, the selection of a physiological indicator should invariably entail the simultaneous consideration of the indicator's biological matrix and the formulation of a realistic sampling design. Unfortunately, the 'secondary elements' also have their own important confounding factors, as well as technical, practical and operational limitations and requirements that must be known and controlled (Palme 2019).

Overall, both positive and negative characteristics of all elements involved (e.g. indicator, matrix, species, sampling method) need to be identified and considered together when physiological indicators are intended to be used in welfare assessments (Romero and Beattie 2022). When designing the physiological indicator's sampling protocol, factors such as the number of samplings possible, the collection method (refined for protecting the animal's welfare, avoiding stress of collection or other altered results and ensuring staff safety) and the storage possibilities (some matrices are more stable than others) should also influence the final selection.

Table 2. Examples of considerations when identifying adequate physiological indicators of welfare

Element	Examples of considerations
Physiological indicator	Has it been validated for the target species and context? Which welfare-related aspect is the indicator related to and how? How does the indicator respond across the entire range of welfare status (e.g. it may be valid for detecting differences between intermediate and poor welfare conditions but not between intermediate and very good conditions)? What is the welfare sensitivity of the indicator? What are its confounding factors and how can they be controlled? Is its quantification feasible (e.g. laboratory of reference, validated analytical methods, sufficient budget)?
Matrix	Has it been validated for the target indicator, species and context? What specific timeframe is the matrix (and its sampling design) providing information about? How is the matrix affecting the indicator's sensitivity? What are its confounding factors and how can they be controlled? What is the right sampling design considering the target animal, the context and the welfare assessment aim (e.g. number of samples, sampling frequency, collection method)? What are its sampling requirements, limitations and risks? What are its storage requirements? What are its analytical requirements? Have the laboratory indicator's extraction procedures been validated for this matrix?
Species	Have the indicator and the matrix been validated to assess the welfare of this species? What are its confounding factors and how can they be controlled (e.g. chronobiological changes, indicators' metabolisation and excretion routes, physiological differences between sexes)? What are its common or more probable welfare problems in zoological institutions?
Individual and context	How are the physiological indicator and matrix affected by the biological status of the individual (e.g. age, sex, physiological and health status, pregnancy)? What is the expected welfare status of the individual? What is the welfare assessment aim in relation to the individual? What contextual information is needed to correctly interpret the physiological indicator? What sampling procedures are applicable to the individual? Can the individual meet all sampling requirements (e.g. minimum amount of sample, specific location)? What is the management of the animal during the sampling (trained or desensitised versus sedated or anaesthetised versus forced) and how could it affect the physiological indicator (e.g. voluntary and calm versus forced and stressed)? What is the balance between the potential benefits (welfare-related information gain) and problems (e.g. animal stress, health risk) related to the use of this indicator, matrix and sample protocol in the individual?

A three-step process is suggested to facilitate an adequate selection of physiological indicators for welfare assessments:

1. Establish the welfare assessment aim: set the target individual(s), the timeframe intended for monitoring, the expected range of welfare of the target animal(s) and the potential welfare problems or threats to identify.
2. Know the context and define the working framework: consider both the species-specific characteristics and individual particularities, the possibility of collecting the necessary contextual information, animal management and training possibilities related to sampling and the budget
3. List and describe the most adequate indicators: consider their connection to welfare, their complementarity with other welfare indicators, the most adequate biological matrix and sampling method for the timeframe monitored and expected welfare status, the possibility to know or control all confounding factors, the degree of validation of the indicators and the feasibility of the laboratory analysis.

For instance, the selection of physiological indicators for welfare assessment would vary depending on the species (and the matrices available), whether the aim is to assess the welfare status of an individual or group, whether the samples can be individualised and identified or not and whether the studied context is more specific (e.g. the effect of an environmental enrichment action) or more general (e.g. adaptation to a new facility and management), among many other considerations.

Challenges and limitations

The challenges of selecting appropriate physiological indicators described previously are exacerbated by the lack of validated indicators for most species kept in zoos (Campbell-Ward 2023) and by the complex analytical requirements. Studies describing the validity, reliability and feasibility of physiological indicators for wild species are necessary, as well as for each potential matrix of origin and species. Furthermore, the validation of the collection, storage procedures and laboratory work involved (ranging from the processes to extract the indicator from its matrix of origin to the quantification method) is as essential as the biological validation in pursuit of reliable results (Buchanan and Goldsmith 2004; Palme 2019). Fortunately, there are important validation steps, such as setting the minimum amount of sample necessary to perform adequate analysis, studying the stability of matrices under certain storage conditions or identifying and describing the influence of confounding factors associated with the matrix (such as body location, type and colour, sources of contamination) which can be relatively easily undertaken in zoological institutions.

As mentioned, the possibility of performing adequate laboratory treatment and analysis of the samples is critical. This implies that all samples should be processed and analysed uniformly and in the same laboratory (Palme 2019; Schoenemann and Bonier 2018). Interestingly, when the laboratory personnel know the welfare monitoring objectives, sampling design and context, the expected errors associated with sample treatment and analysis can be strategically managed to reduce or avoid affecting biological interpretation of the indicator. For instance, an adequate distribution of samples in several enzyme-linked immunosorbent assays (ELISAs) can prevent intraassay variability errors from being attributed to different individuals, conditions or times. Integrating laboratory personnel into the entire welfare assessment seems an advantageous strategy. Staff expertise and experience working with the indicator, matrix and species, as well as the existence of validated laboratory protocols or the availability of adequate equipment among others are also relevant factors that condition the potential use of physiological indicators and the quality of their interpretation when assessing animal welfare (Schoenemann and Bonier 2018).

Conclusion

Achieving the highest standards of animal welfare in zoos is essential, grounded in ethical considerations and vital for fulfilling the educational and conservation roles of modern zoos. Evidence-based approaches are crucial in assessing and monitoring animal welfare and the use of both behavioural and physiological indicators can be useful. Behavioural indicators, widely recognised as crucial, demand careful selection, development and validation. While behavioural observations provide valuable insights, challenges such as observer bias and external factors need consideration. The significance of physiological indicators should be emphasised, acknowledging their diversity and complexity. The careful selection, validation and interpretation of physiological indicators, in conjunction with behavioural assessments and other welfare indicators, offer a holistic approach to welfare evaluation in zoo-housed animals.

References

Altmann J. (1974) Observational study of behavior: Sampling methods. *Behaviour* 49(3–4): 227–267. doi:10.1163/156853974X00534

Binding S., Farmer H., Krusin L., Cronin K. (2020) Status of animal welfare research in zoos and aquariums: Where are we, where to next? *Journal of Zoo and Aquarium Research* 8(3): 166–174. doi:10.19227/jzar.v8i3.505

Botreau R., Veissier I., Butterworth A., Bracke M.B.M., Keeling L.J. (2007) Definition of criteria for overall assessment of animal welfare. *Animal Welfare* 16(2): 225–228. doi:10.1017/S0962728600031390

Bracke M.B.M., Hopster H. (2006) Assessing the importance of natural behavior for animal welfare. *Journal of Agricultural and Environmental Ethics* 19: 77–89. doi:10.1007/s10806-005-4493-7

Brambell F.W.R. (1965) R*eport of the Technical Committee to Enquire into the Welfare of Animals Kept Under Intensive Livestock Husbandry Systems.* London, UK: Her Majesty's Stationery Office.

Broom D.M., Johnson K.G. (2019) Stress and animal welfare. In: Broom D.M., Johnson K.G. (eds.). *Stress and Animal Welfare* (Vol. 19). Springer International Publishing.

Browning H. (2023) Welfare comparisons within and across species. *Philosophical Studies* 180(2): 529–551.

Buchanan K.L., Goldsmith A.R. (2004) Noninvasive endocrine data for behavioural studies: The importance of validation. *Animal Behaviour* 67(1): 183–185. doi:10.1016/j.anbehav.2003.09.002

Campbell-Ward M. (2023) Quality-of-life assessments in zoo animals: Not just for the aged and charismatic. *Animals* 13(21): 3394. doi:10.3390/ani13213394

Clegg I., Borger-Turner J., Eskelinen H. (2015) C-Well: The development of a welfare assessment index for captive bottlenose dolphins (*Tursiops truncatus*). *Animal Welfare* 24(3): 267–282. doi:10.7120/09627286.24.3.267

Diana A., Salas M., Pereboom Z., Mendl M., Norton T. (2021) A systematic review of the use of technology to monitor welfare in zoo animals: Is there space for improvement? *Animals* 11(11): 3048. doi:10.3390/ani11113048

Dickens M.J., Romero L.M. (2013) A consensus endocrine profile for chronically stressed wild animals does not exist. *General and Comparative Endocrinology* 191: 177–189.

Gormally B.M.G., Romero L.M. (2020) What are you actually measuring? A review of techniques that integrate the stress response on distinct time-scales. *Functional Ecology* 34(10): 2030–2044. doi:10.1111/1365-2435.13648

Hill S.P., Broom D.M. (2009) Measuring zoo animal welfare: Theory and practice. *Zoo Biology* 28(6): 531–544. doi:10.1002/zoo.20276

Hosey G., Melfi V., Pankhurst S. (2009) Behaviour. In: Zoo Animals: Behaviour, Management and Welfare. Oxford, UK: Oxford University Press, 82–128.

Lawrence A.B., Vigors B., Sandøe P. (2019) What is so positive about positive animal welfare?—A critical review of the literature. *Animals* 9(10): 783. doi:10.3390/ani9100783

Le Neindre P., Bernard E., Boissy A., Boivin X., Calandreau L., Delon N., Deputte B., Desmoulin-Canselier S., Dunier M., Faivre N., Giurfa M., Guichet J.L., Lansade L., Larrère R., Mormède P., Prunet P., Schaal B., Servière J., Terlouw C. (2017) Animal consciousness. *EFSA Supporting Publications* 14(4): 1196E. doi:10.2903/sp.efsa.2017.EN-1196

Manteca X., Amat M., Salas M., Temple D. (2016) Animal-based indicators to assess welfare in zoo animals. *CABI Reviews* 11(010): 1–10. doi:10.1079/PAVSNNR201611010

Mellor D.J. (2016) Updating animal welfare thinking: Moving beyond the "Five Freedoms" towards "A life Worth Living". *Animals* 6(3): 21. doi:10.3390/ani6030021

Palme R. (2019) Non-invasive measurement of glucocorticoids: Advances and problems. *Physiology and Behavior* 199: 229–243. doi:10.1016/j.physbeh.2018.11.021

Ralph C.R., Tilbrook A.J. (2016) The usefulness of measuring glucocorticoids for assessing animal welfare. *Journal of Animal Science* 94(2): 457–470. doi:10.2527/jas.2015-9645

Romero L.M., Beattie U.K. (2022) Common myths of glucocorticoid function in ecology and conservation. *Journal of Experimental Zoology Part A: Ecological and Integrative Physiology* 337(1): 7–14. doi:10.1002/jez.2459

Sadoul B., Geffrey B. (2019) Measuring cortisol, the major stress hormone in fishes. *Journal of Fish Biology* 94(4): 540–555. doi:10.1111/jfb.13904

Salas M., Manteca X., Abáigar T., Delclaux M., Enseñat C., Martínez-Nevado E., Quevedo M.A., Fernández-Bellón H. (2018) Using farm animal welfare protocols as a basis to assess the welfare of wild animals in captivity—Case study: Dorcas gazelles (*Gazella dorcas*). *Animals* 8(7): 111. doi:10.3390/ani8070111

Schoenemann K.L., Bonier F. (2018) Repeatability of glucocorticoid hormones in vertebrates: A meta-analysis. *PeerJ* 6: e4398. doi:10.7717/peerj.4398

Sherwen S.L., Hemsworth L.M., Beausoleil N.J., Embury A., Mellor D.J. (2018) An animal welfare risk assessment process for zoos. *Animals* 8(8): 130. doi:10.3390/ani8080130

Staley M., Conners M.G., Hall K., Miller L.J. (2018) Linking stress and immunity: Immunoglobulin A as a non-invasive physiological biomarker in animal welfare studies. *Hormones and Behavior* 102: 55–68. doi:10.1016/j.yhbeh.2018.04.011

Tallo-Parra O., Salas M., Manteca X. (2023) Zoo animal welfare assessment: Where do we stand? *Animals* 13(12): 1966. doi:10.3390/ani13121966

Tarlow E.M., Blumstein D.T. (2007) Evaluating methods to quantify anthropogenic stressors on wild animals. A*pplied Animal Behaviour Science* 102(3–4): 429–451. doi:10.1016/j.applanim.2006.05.040

Veasey J.S., Waran N.K., Young R.J. (1996) On comparing the behaviour of zoo housed animals with wild conspecifics as a welfare indicator, using the giraffe (*Giraffa camelopardalis*) as a model. *Animal Welfare* 5(2): 139–153. doi:10.1017/S0962728600018571

Whitham J.C., Wielebnowski N. (2013) New directions for zoo animal welfare science. *Applied Animal Behaviour Science* 147(3–4): 247–260. doi:10.1016/j.applanim.2013.02.004

Yon L., Williams E., Harvey N.D., Asher L. (2019) Development of a behavioural welfare assessment tool for routine use with captive elephants. *PLoS ONE* 14(2): e0210783. doi:10.1371/journal.pone.0210783