Inspiring systematic inclusion of individual animal states to enhance the quality of research

Janire Castellano Bueno [1,2,] * Vittoria Elliott [1,3]

[1] Wild Animal Initiative, Minneapolis, MN 55437, USA
[2] Newcastle University, Newcastle upon Tyne, UK
[3] National Museum of Natural History, Smithsonian Institution, Washington, USA

* Corresponding author: janire.castellanobueno@gmail.com

Abstract

Studies on animals continue to attract criticism over data quality, reproducibility and generality of findings, yet one source of variation remains rarely addressed: differences in individuals' affective states. In this paper, we suggest that evaluating affect should be considered standard good practice in ecological and behavioural research with wild animals, alongside familiar variables such as sex, age or dominance rank. Affective states, viewed here as integrated products of animals' past and current experiences, influence how individuals perceive and respond to experimental treatments, environmental conditions and human activities, and can systematically alter behavioural, physiological and ecological outcomes.

We first place affective states within a review of existing discussions about replication, bias and rigour in animal research, and argue that unaccounted for differences in affect can generate unexplained variation, apparent outliers, selective exclusion of individuals and difficulties in reproducing results across sites or studies. By reviewing relevant animal welfare science research, we show how behavioural, somatic, physiological and cognitive indicators can be used to evaluate welfare and infer affective states in both captive and free-ranging animals, increasingly with minimally invasive or remote methods. Building on this, we outline practical ways that differences in affect can be inferred by incorporating welfare assessments into wildlife studies. By systematically collecting welfare-relevant data throughout a project, affective status can be used as an explanatory or control variable in statistical models, to detect sampling biases, and to interpret behavioural patterns.

Finally, we discuss ethical and logistical implications. Integrating welfare assessments has the potential to support the 3Rs by reducing unnecessary repetition, increasing the amount learned from each animal and helping to identify welfare problems earlier. Despite challenges such as limited species-specific indicators

and added analytical complexity, routinely accounting for affective states is a feasible, important step toward more robust, informative, and ethically defensible research with wild animals.

Keywords: Reliability, Reproducibility, Rigor, Affective State. Animal Welfare, Behavioural Ecology

1. Introduction

With this paper, we aim to illustrate how considering animal affective* states (a key component of welfare*), which are known to influence behavioural responses (Montalcini et al., 2025), can enhance the quality (consistency, accuracy) and utility of data collected. Individual variation in internal welfare states (i.e.,"affective state") can impact the consistency of individual animal responses, and thus study results, influencing analysis and interpretation. We hope that by raising awareness among the scientific community working with wild animals of how differences in an individual's affective states can lead to variation in how study subjects respond to treatments or their environments, researchers will consider including ways to assess affect in future studies. Our goal is to convince readers that when working with wildlife, best research practice should include collection of data on welfare indicators, so that individual differences in affective state can be accounted for in analysis and interpretation of results.

Specifically, we would like to encourage all those working with animals, to (1) consistently incorporate the collection of data that allows them to evaluate affective states in study designs, (2) consider differences in affective states among subjects and evaluate their impact on study results during analysis and interpretation by applying lessons learned from captive animal welfare science to studies of animals in the wild.

2. Review of data quality issues in science

Recent introspection within the research community has highlighted systemic biases* and data quality* issues in scientific research (Florczak, 2022; Wicherts, 2017). These issues range from concerns about the way that data are collected (e.g., which animals are included in studies and how subjects are selected), and how analysis is conducted (e.g., aggregation and smoothing of variation, selective removal of data such asp-hacking*), to what results are reported (e.g., underreporting of negative results) (Cheleuitte-Nieves & Lipman, 2019; Filazzola & Cahill Jr, 2021; Jilka, 2016; G. D. Smith & Ebrahim, 2002). In several disciplines, such biases have been shown to shape research practice, the pattern of published findings, and the way science is communicated (Bartoš et al., 2024; Bruns et al., 2024). Such practices can lead to potential challenges for replication* of studies, reproducibility* of results, hinder transparency* about research methods used, and can even reduce the generalisability* of results (Cheleuitte-Nieves & Lipman, 2019; Jilka, 2016; Schloss, 2018). In particular, replication and reproducibility have emerged as consistent issues across various academic disciplines, with areas such as psychology and biomedical research exhibiting alarmingly low reproducibility rates*; 11% to 82% (Begley & Ellis, 2012; Freedman et al., 2015; Mundinger et al., 2025).

Such trends could even call into question the reliability of established knowledge (Guttinger, 2020) and raise fundamental questions about research methodologies, potentially contributing to scepticism in both scientific communities (Baker, 2016) and public perception (Anvari & Lakens, 2019).

2.1 Data quality issues in animal research

Although, to our knowledge, the issues of replication and reproducibility have not been comprehensively assessed across animal science research, it is reasonable to expect that similar concerns could arise for studies within ecology, animal behaviour, cognition, and welfare science (Brecht et al., 2021; Cheleuitte-Nieves & Lipman, 2019; Fidler et al., 2017; Mundinger et al., 2025; Parker et al., 2016). Indeed, Brecht et al., (2021) and Beran (2020) both recently discussed the importance of replication for animal behaviour and cognition studies and highlighted that these disciplines appear to have low replication rates. Both studies stressed the need for greater methodological rigour*, including standardized experimental procedures, appropriate statistical power, proper control conditions, and systematic data collection protocols that minimize measurement error and researcher bias.

Some recent studies have identified particular barriers within animal research disciplines that can hinder replication of studies and contribute to data reporting issues*. For example, Shaw et al., (2021) and Filazzola & Cahill Jr, (2021), focusing on endangered species and field ecology respectively, both found that resource limitations and regulatory constraints can prevent studies from being replicated (resource-level constraints), while biases against the publication of negative results can affect the accurate reporting of findings (publication-level bias). Similarly, Nawroth & Gygax, (2020), focusing on farmed animal welfare science, emphasised the need for more replication efforts to overcome differences in facilities and increase generalisability of results across heterogeneous contexts.

A number of factors have been highlighted as contributing to data quality issues when they are not appropriately accounted for in studies (Voelkl et al., 2020). For example, study results are subject to the impacts of individual variability (i.e., genetics, age, sex differences), environmental influences (i.e., housing, diet, habitat), methodological approaches (e.g., observer bias, capture, observer disturbance effects, handling), and differences between facilities, equipment, or context (e.g., lack of standardisation across laboratories, differences in equipment use) (Cheleuitte-Nieves & Lipman, 2019; Jilka, 2016). Failure to adequately consider these factors has been implicated in introducing biases and leading to misinterpretations that could potentially present challenges for replication and reproducibility of animal studies (Voekl et al., 2020). For instance, if studies do not control for or adequately consider the impacts of sex differences (i.e., assess, record and include sex as a variable in analysis) on their research question, they run the risk of wrongly attributing behavioural observations associated with sex differences to some other variable or environmental factors being studied (Das, 2002).

When differences between study subjects and conditions are inappropriately considered or unaccounted for, these issues can impact the relevance, utility, and generalisability of research findings and can contribute to challenges for replication. For example, relevance is affected when uncontrolled variables distort the link between the findings and the biological question; utility is reduced when results, derived from uncontrolled or inconsistent conditions, cannot inform future research or applications; and generalisability is limited when findings only hold under specific, conditions (such as individual affective states). These limitations make it difficult for other researchers to reproduce the results, ultimately undermining confidence in the study's conclusions.

2.2 What has been done to combat these issues

Various authors have aimed to address replication and reproducibility issues in animal studies (Brecht et al., 2021; Volsche et al., 2023), suggesting several approaches that might increase replication and reproducibility, including systematic heterogenisation* (i.e., inducing controlled variability into the study design) (Richter, 2017; Richter et al., 2009; Voelkl et al., 2020), preregistration or registered reports (Chambers & Tzavella, 2022), multi-laboratory designs (Voelkl et al., 2018), sharing of databases (e.g., ManyBirds, ManyPrimates), and inclusion of more detailed subject information (Volsche et al., 2023). A number of guidelines have also been published that make recommendations for how to decrease potential biases (ARRIVE (Percie Du Sert et al., 2020), PREPARE (Smith et al., 2018). Many of the suggested approaches and guidelines highlight understanding and accounting for individual animal context differences as key to reducing biases, however, such discussions of data bias, reproducibility, and other proposed solutions frequently overlook differences in affective state (i.e., as indicated by the assessment of welfare status*), among individual study animals, as a source of variation that might need accounting for (Volsche et al., 2023; Yeates, 2024).

Although no guidelines explicitly advocate for the routine inclusion of welfare assessments in animal studies, some frameworks do implicitly recommend considering animal affective states by encouraging researchers to consider factors relating to it. For instance, the STRANGE framework emphasises the need to consider the specific characteristics of the animals used for research, such as their social background, trappability and self-selection, rearing history, acclimation and habituation, natural changes in responsiveness, genetic makeup, and past experiences (Webster & Rutz, 2020), all of which can contribute to differences in affective state.

3. Affective states and their impact on replication and data quality in animal studies

3.1 Affective states

Affective states are defined in terms of an individual's subjective experiences*, that is, a sentient* individual animal's accumulated emotions and feelings in response to their circumstances (internal states and environmental conditions) (Marchant-Forde, 2015). In this paper, we consider affective state as a critical component of animal welfare that influences how individuals respond to their circumstances.

Affective states can serve as an aggregate measure of past and present experiences. Thus, they not only reflect the contemporary state of an individual animal but also reflect their past experiences and their responses to them (Mellor, 2016; Veissier et al., 2012). Recent frameworks emphasize this temporal integration, situating animal welfare as an ongoing reflection of both immediate and cumulative affective experiences. Evaluating affective states can therefore be used to consider not only the current impacts of factors such as environmental conditions, recent events, and animal handling, but can also indicate the unique integrated affective status of individuals.

Affect cannot be directly measured, but it can be inferred from various somatic, behavioural, and cognitive indicators that reflect physical, mental and emotional experiences. Individual affective states can thus be inferred by assessing 'welfare indicators' that demonstrate physical, somatic, physiological, mental and cognitive components of an individual's welfare. Evaluating multiple indicators that measure different aspects of the welfare of an individual animal, therefore, provides holistic animal-centric insight into how they are experiencing their circumstances. It indicates the individual interpretation of the sensory inputs they receive from their physical condition, internal state, and their environment (Browning, 2023; Mellor, 2015). In practice, affect is inferred by assessing a combination of welfare indicators, such as postures (MacLellan et al., 2022), alertness (Harvey et al., 2023a), cognitive bias (Bethell, 2015; Crump et al., 2018; Lee et al., 2016; Monk et al., 2020), heart rate (Kitajima et al., 2021; von Borell et al., 2007), and a variety of behaviours (Browning, 2023; Browning et al., 2024) among others.

3.2 Why include affective states in study frameworks

An animal's experiences not only impact their affective state (i.e. their internal experiences), but also how they behave (i.e., how they respond) (Carranza-Pinedo et al., 2025). It follows, therefore, that the behaviours that researchers observe and seek to understand, are shaped by the internal state of the individual animal, and their unique affective experiences in the context of their environment, including both natural and anthropogenic factors (Loss et al., 2021). Despite this well-established link between affective state and behaviour (Dawkins, 2004; Watters et al., 2021), research on wild animals often overlooks variation in individual affective states, as well as the origin of this variation (Volsche et al., 2023; Yeates, 2024).

While not yet widely accepted, there is increasing recognition within laboratory settings that an individual subject's affective states play a critical role in the quality of the data collected from both experimental and observational studies (Poole, 1997). There is growing evidence that variations in housing conditions, handling practices and experimental conditions that impact animal welfare, introduce confounding variables into data gathered from animal subjects (Brill et al., 2019; Poole, 1997). Studies showing impacts of early life experiences on later health and welfare reinforce the importance of considering the impacts of pre-study animal welfare status on experimental or observational study outcomes (Kinnally et al., 2019). Animal welfare science and experimental laboratory studies have discussed the influence of affective states of study subjects on data quality for animal behaviour and cognition studies (Baciadonna et al., 2024; Cassidy et al., 2020; Descovich et al., 2019; Truelove et al., 2020) suggesting that positive affective states increase model validity, reproducibility and reliability of data (Grimm, 2018; Poole, 1997; Prescott & Lidster, 2017). For example, studies have shown that not only do animals in less stressful housing conditions produce less variable data, but so do animal subjects that experience less early-life stress (Guerrero-Martin et al., 2021). These findings suggest that impacts of affective state on data quality should not only be considered with regard to influences arising from the laboratory environment and study procedures such as handling, instrumentation and observer presence, but also with respect to the effects of variation in environmental conditions, whether anthropogenic or natural.

Because differences in the affective status of study subjects can significantly influence how and how much individuals respond to their circumstances, including but not limited to how they respond to experimental treatments, failure to record indicators of such differences can result in misleading interpretations of observed behaviours (Cassidy et al., 2020; Grimm, 2018; Poole, 1997; Prescott & Lidster, 2017).

Variation in affective state among individual subjects can also manifest as outliers in the data, which may be excluded from further analysis. For example, variation in responses related to differences in affective state may be treated as statistical noise or as anomalies that are then excluded from analysis, and interpretation (Bushby et al., 2018; Hecht, 2021). Individual subjects may also be removed based on injury or other welfare concerns as part of initial exclusion criteria (ARRIVE). Such exclusions are rarely reported, but can skew the data collected (Percie du Sert et al., 2019). Consequently, overlooking differences in affect between subjects can potentially introduce challenges for replicating experimental results in different environments or with different individuals and prevent opportunities for drawing robust conclusions across contexts (Prescott & Lidster, 2017; Truelove et al., 2020). Thus, failure to include welfare assessment in animal studies may be contributing to replication challenges and may prevent generalisation.

Furthermore, although it may not necessarily be the principal subject of study, animal behaviours are often monitored in other research involving animal subjects and are used as a response variable to indicate a reaction to a given experimental treatment, stimulus or circumstance. Understanding and correctly interpreting the impacts of affective state on animal behaviour is, therefore, critical for obtaining accurate results and inferring causation not only in studies of animal welfare or animal behaviour, but in all research involving animals.

While laboratory and procedural impacts on affective states are increasingly being considered, natural variation in affective states, stemming from animal history, life events, or previous experiences, is still often overlooked. Without recognising and accounting for both types of variation, accurate interpretation of observed behaviours remains challenging.

3.3 Progress made in animal welfare research

3.3.1 Improvements in laboratory studies

Increasing understanding of captive animal experiences and the ethical implications of poor husbandry practices for animal welfare has led to greater awareness of animal affective states and the need to consider differences in individual animal contexts, including variations in affective state more effectively in research with animal subjects. This has given rise to the development of best-practices for animal use, handling, and engagement. Animal welfare science research has aided understanding and helped generate a variety of guidelines, recommendations and assessment protocols to facilitate good and ethical practice for interactions with animals, including husbandry practices for farmed animals, interactions with companion animals and care and handling of animal subjects in research (Hawkins et al., 2011).

These guidelines and regulatory procedures requiring welfare evaluations are now a standard component of most research involving direct interactions with animals, particularly those under human care. Oversight bodies such as Institutional Animal Care and Use Committees (IACUCs) in the United States, the Home Office in the United Kingdom, and ethics review boards mandated under the European Union Directive (*Directive 2010/63/EU.*) ensure that animal use in research is subject to rigorous ethical and scientific scrutiny. These committees evaluate both the necessity of proposed studies, requiring researchers to demonstrate that the potential benefits justify any impact on animal welfare. Central to these frameworks is adherence to the 3Rs principle (Replacement, Reduction, and Refinement) which encourages efforts to replace the use of animals wherever possible, reduce the number required for valid results, and refine procedures to minimize pain, stress, or lasting harm.

The implementation of these frameworks relies on the tools and evidence provided by Animal Welfare Science (AWS). By supplying standardized indicators across physical, physiological, behavioural and cognitive domains through which affective states and other components of welfare status can be inferred, AWS makes it possible to translate ethical principles into practice. In doing so, it supports the 3Rs and ensures that oversight is guided by empirical assessment, thereby reinforcing the legitimacy and effectiveness of including welfare evaluations in animal research and emphasizing the importance of understanding differences in affective state of animal subjects.

Although, by following such guidelines, interactions and exposure to potential stressors can be carefully managed in captive settings, achieving fully homogeneous welfare states among all individuals is not always possible or even desirable. Inherent differences in temperament, history, and other intrinsic characteristics might make true homogeneity unattainable. Efforts to make welfare completely uniform may not always be appropriate, as enforcing uniformity could have negative effects on other populations or environmental conditions. In addition, preserving natural variation within the dataset makes it more likely that the results will reflect the full range of outcomes seen in animals. This increases the generalisability of the findings, which is important for the broader application of animal research(MacLellan & Lalu, 2024). Therefore, it is important to monitor and assess welfare indicators that demonstrate individual differences in affective states, even in controlled settings, to understand their influence on study outcomes and to support accurate interpretation of results.

3.3.2 Improvements in field studies

There is growing recognition within conservation science of the relevance of understanding natural variation in affective state, as it has been shown that it can directly influence the outcome of conservation programs (e.g., translocation success; Reading et al., 2013; Swaisgood, 2010). Similarly, awareness of the impacts that welfare experiences and associated differences in the affective state of study subjects can have on research results is increasing among wildlife researchers. For instance, studies have shown that data loggers can impact bird behaviours (Eberhart-Hertel et al., 2025; Geldart et al., n.d.), while post-release observations of fish with trackers suggest their behaviours are influenced by handling procedures (LaRochelle et al., 2025; Niella et al., 2023). Researchers are therefore taking steps to minimize and mitigate such impacts (e.g., (Eldegard et al., 2024) and additional resources for applying the 3Rs principles with wildlife are becoming increasingly available (Zemanova, 2021). However, researchers generally only consider affective states and other components of welfare in circumstances where they expect that the methods used to assess animal behaviours are likely to directly impact the welfare of the study animals and could affect the quality of the data collected (i.e., differences caused by the study itself, e.g., capture or instrumentation such as biologging approaches) instead of considering influences of differences in affective state on study outcomes (Beaulieu & Masilkova, 2024).

Most research that includes and evaluates indicators of the welfare of study subjects tends to focus on laboratory conditions or other human-induced impacts of the study itself, such as housing, setup, capture, handling, and treatments (Zemanova, 2021). Natural variations in affective states and other components of welfare are generally less frequently considered. As a result, relatively few studies, particularly those conducted in the wild, incorporate measurements that would allow investigators to account for affect as a factor or variable of interest.

4. Benefits of including affective states

4.1 Affect as a motivator

Affect (the way an individual experiences a given stimulus) drives or motivates how, when, and how much individuals will respond to a given stimulus. Thus, affect motivates individual behavioural responses. The human psychology literature emphasises that it is important to acknowledge how affect influences motivations, and both variation and consistency in responses must be accounted for to form robust conclusions about behavioural observations (Ekkekakis, 2013; Lench et al., 2011; Lerner et al., 2015; Li et al., 2024). Given the increasing awareness of animal emotions and a greater understanding of how to evaluate affective states (Boissy et al., 2007; Panksepp, 2011; Paul et al., 2005), applying these same principles to animal studies would likely also be valuable. Integration of welfare assessment would allow investigators to consider how variation in the affective states of animal subjects is influencing the effects of motivators and how they might also drive differences in animal responses.

Importantly, the affective states of an individual can impact not only their contemporary behaviour, but also future behaviour. Indeed, an animal's past and present experiences contribute to their current affective state

(Veissier et al., 2012) (e.g., thirst, hunger, tiredness), which subsequently drives their motivations to perform certain actions, such as hunting, socialising, resting, or foraging (Beausoleil et al., 2018). Increasing awareness of this association highlights the importance of understanding how animal motivations relate to their affective states, which can help explain variation in behaviour and identify the drivers of responses to stimuli (Mason & Bateson, 2017).

4.2 Affect as a driver of variation

Since differences in affective states can greatly affect the way that individual animals respond to experimental or environmental conditions, there will always be considerable variation in the observed responses (e.g. Mendl et al., 2009; Paul et al., 2005). Thus, affective states can be an important driver of variations in both physiological and behavioural response during observation (Beaulieu, 2024a). Therefore, just as ways to determine sex, age, relatedness, or rank are considered important to include and control for in animal studies across settings, ways to assess welfare indicators and thus evaluate affective differences should also be included (Beaulieu, 2024a; Montalcini et al., 2025).

4.3 Affect to strengthen interpretation

Considering affective status when discussing results could improve the interpretation of behavioural studies (See Figure 1), contextualising observed patterns and increasing interpretive power (Poole, 1997). With an understanding that an animal's affective state motivates their actions and reactions, inclusion of welfare assessment provides valuable insight for interpretation of observations and patterns of behaviour. Not only does it provide insight into variations in behavioural responses, but also potentially into triggers and drivers for certain behaviours such as dispersal, fight-flight responses, and positive and negative interactions with conspecifics (Anderson, 2016; Budaev et al., 2019; Kennedy et al., 2014; Papini et al., 2019; Takahashi, 1986).

4.4 Affect as a control variable

Explicitly supporting the incorporation of affective state as a variable in study design and/or as a control variable during statistical analysis, depending on the type of study (for example experimental vs observational), could provide a framework for researchers to systematically account for its effects on results, and at the very least could provide a key explanatory variable. Assessing indicators of welfare allows researchers to report differences in affect among subjects and transparently note any exclusions, such as due to concerns about their welfare. Doing so improves both the interpretation and replicability of experiments, as future studies can account for these factors when designing analyses.

4.5 Ethical standards

Failing to account for animal affective states in animal studies might not only compromise scientific quality and reproducibility of findings but may sometimes also undermine ethical scientific practice. Integrating affective status into research is also likely to help minimise harm and optimise resource allocation by encouraging the systematic identification of welfare-related issues.

Animal welfare legislation, ethical review processes, and the 3Rs framework (Replacement, Reduction, Refinement) emphasise that harm to animals must be minimised and that studies should yield meaningful knowledge that justifies their use (MacArthur Clark & Sun, 2020; Russell & Burch, 1959). If affective states are not adequately monitored and reported, animals may suffer unnecessarily, which conflicts with ethical standards of refinement and transparency. Furthermore, poor welfare (resulting in negative affective states) can alter behaviour, physiology, and cognition in ways that bias results, leading to unreliable or non-generalisable data (D. J. Mellor, 2016). This means that failing to integrate assessment of welfare risks both ethical non-compliance and wasted research efforts.

Conversely, explicitly considering affective states in experimental design and reporting helps to minimise harm by identifying and mitigating sources of stress or suffering, thereby aligning with refinement principles.

Although originally developed in laboratory contexts, the 3Rs can help guide ethical practice in the wild, for example by promoting non-invasive observation in place of capture (Replacement), improving study design to reduce sample sizes (Reduction), and refining capture, tagging, or handling methods to minimize stress (Refinement). Extending the 3Rs to fieldwork ensures that the pursuit of ecological and conservation knowledge is aligned with animal welfare considerations.

A significant consequence of poor reproducibility in animal studies is the unnecessary use of animals in follow-up research. When initial findings are difficult to replicate due to unassessed and therefore unobserved differences in affect, researchers may need to repeat experiments or increase sample size, subjecting additional animals to capture, handling, or potentially distressing exposure to humans. This repeated exposure and distress raises ethical concerns, directly conflicting with the principle of minimising harm in scientific studies and with the "Reduce" principle of the 3Rs framework (Prescott & Lidster, 2017).

Beyond ethical considerations, poor welfare standards may have substantial resource implications. Collection of welfare measures can optimise resource allocation by ensuring that data collected are not compromised by unrecognised differences in affect reducing the likelihood that studies will need to be repeated due to poor validity (Franco et al., 2014). Addressing these inefficiencies may help conserve valuable time, funding, and personnel effort, freeing up resources to support other important research initiatives and accelerating progress across the field.

4.6 Differences in affective states in the wild

The welfare of wild animals is increasingly recognised as a vital yet complex area of ethical and scientific inquiry (Beausoleil & Mellor, 2015; Browning & Veit, 2023; Sleegers et al., 2025). While some debate its conceptual clarity and practical feasibility, there is growing support suggesting that consideration of affective states in wild populations is both relevant and necessary (Beaulieu, 2024a; Beausoleil et al., 2018; Browning, 2022; Browning et al., 2024).

The benefits outlined above are just as relevant, if not more so, for studying wild populations. Considering the extent of research involving wild animals, and an increasing understanding of the impacts of research activities on the affective states of wild populations, integrating welfare assessments into wildlife studies would seem essential. Furthermore, interindividual differences in affective states are likely to be exacerbated in the wild where study conditions can vary considerably and individual experiences are highly heterogeneous. The issues associated with not considering affective differences among individuals could likewise also be exacerbated in ecological or animal studies conducted in the wild. Greater mobility of free-ranging animals (Williams et al., 2020), variable spatial and temporal scales (Powers & Hampton, 2019), and the limited ability of investigators to have access to individual information, to consider a sufficient number of replicates or to control environmental conditions (Filazzola & Cahill Jr, 2021) can increase unexplained variation in the study results. Despite potential logistic challenges, animal studies conducted in the wild should, nevertheless, be able to yield replicable findings, provided that studies adhere to standardised protocols and conduct welfare assessments to contextualize their results (Nakagawa & Parker, 2015). Although homogenising life-experiences and individual affective states might be neither possible nor desirable in the wild, considering affect as a potential control variable and collecting information on differences in affective states is likely to facilitate more accurate interpretation of the study results.

Some argue that, since in the wild affective status is not shaped by rearing practices or housing conditions, subjects naturally represent the full range of variation (Niśkiewicz et al., 2024; Schad et al., 2025). It has also been suggested that standard randomized sampling methods used for studying animals in the wild can account for differences in study subjects (Gregol et al., 2024). Whilst possible, random samples are unlikely to represent the full range of naturally occurring variation in affective states. For instance, stressor effects limited to some subjects may impact behaviours such that some individuals do not appear within the study frame. Animals that have particular affective states (such as lethargy due to dehydration) or that have had certain past affective experiences (such as fear-inducing events) may also 'opt-out' of studies. For example, 'self-selection' through avoidance of traps (Sockman & Beaulieu, 2023), camera-shy behaviours, reduced mobility, and other reclusive characteristics, can all remove components of the natural variation and contribute to limitations in the range of affective states of the studied subjects (Cook et al., 2021). This unintended sampling effect could skew the study towards investigating only certain groups or individuals. For example, the study might only capture those individuals that are too exhausted to move away from the study area. Sample sizes may also be too small to capture the range of affective differences, while statistical removal of outliers might exclude individuals with particular affective states. Such effects of sampling methods are likely to bias or skew results from being representative of the natural range of physical and affective states. Importantly, incorporating relevant welfare assessment tools into studies that allow investigators to determine affective differences between study subjects makes them aware of and able to account for such factors in their interpretations.

4.7 Overall benefits of monitoring affect for rigor and repeatability

Using welfare assessments to collect data about affect can help inform studies by providing valuable information about the differences between study subjects. This improves researchers' ability to draw informed conclusions and formulate valid interpretations, making studies more rigorous. In doing so, welfare assessments facilitate an understanding of how generalisable study results are and provide insight into factors associated with differences between study subjects that might influence repeatability and generalization.

Table 1. Summary of potential benefits and challenges of incorporating affective states into animal studies.

Potential benefits
Incorporating the assessment of indicators of affective states into animal studies could help improve the overall robustness and interpretability of research findings by: ● Providing insights about how generalisable findings are likely to be to other populations ● Characterizing key attributes that contribute to reproducibility of results across studies (contexts and timeframe) ● Increasing awareness of sampling effects and skew in representativeness of results ● Supporting more accurate interpretations of behavioural data

- Reducing need for sample exclusion of outliers based on welfare-related factors

Contributing to:

- More efficient use of resources by minimising unnecessary repetition
- Reduction in unexplained variation (data noise), leading to improved understanding of patterns
- Improved welfare for study subjects

Likely challenges

- Risk of model overfitting due to complex datasets, requiring larger sample sizes
- Insufficient species-specific, validated, standardised welfare indicators
- Necessity for multidisciplinary teams to include animal welfare science experts
- Increased time and effort required for study design, implementation, and analyses
- Constraints related to funding
- Logistical challenges in incorporating animal welfare assessments

5. Integrating affective status within studies of wild animals

To help account for affective differences between study subjects, we encourage all investigators working with animals to include more collection of data that can inform an understanding of affective differences in natural study populations.

If all researchers working with wild animals made use of welfare assessment approaches to gather at least some relevant data on affective states of wild animal subjects, it would help build a clearer understanding of the range and variation in individual affective states that naturally occur in the wild. Increased availability of such data would also inform study designs and improve methods, facilitating additional development, simplification, and systematic inclusion of welfare assessments. As a result, investigators would also be able to establish what 'normal ranges' in affective states might be for different species and life stages for future comparison.

To implement such an approach would involve collecting more relevant data about physical, cognitive , and behavioural differences among subjects that can be used to infer and thus evaluate differences in affective states between individuals. To facilitate following good-practices when working with animals in the wild, researchers studying wildlife can incorporate key lessons from lab studies, such as accounting for individual differences in analysis and using validated welfare indicators.

5.1 How monitoring affective status can be integrated into wildlife studies using welfare assessments

The most effective approach for integrating monitoring of affective status into studies of wildlife would be to systematically include several methods of welfare assessment from which affective states of all study subjects can be inferred at various stages of study. Collecting data about *a priori* subject states would establish an understanding of individual affective differences from the outset. By monitoring welfare indicators relevant to affect at various stages of study implementation and following interactions, researchers could establish affective state changes over time, in response to changes in environment or other variables, and offer the opportunity to identify potential impacts of the study itself. Recording physical, cognitive, and behavioural indicators and maintaining records on the inferred affective states of individuals throughout the study and collecting such data for comparison at the inception of new studies would help identify similarities and differences in affective states between individuals, study populations or time periods.

Systemically collecting relevant data would help researchers identify any selection bias in the sampling related to differences in animal states; would serve as a variable informing statistical analyses and would provide additional insight during interpretation of results.

By collecting data on individual animal status investigators could better determine whether there has been selection bias, either intentional by researchers including “healthier-looking” animals, or unintentional “self-selection” through more reclusive individuals remaining unseen, or skew in which animals are captured. Such data also provides insight to the unique characteristics of each study subject. For instance, evaluating indicators of affective states could provide insight to differences between subjects with recent stress levels or differences in early life. Based on our understanding of animals in the lab and how their responses relate to both recent stressful experiences and early-life events, this could help explain observed individual differences in behavioural responses to stimuli, conditions, situations, or other environmental factors being investigated.

In addition to helping researchers understand the unique status and characteristics of their study subjects, data on affective states can be used as an explanatory variable or variables for statistical analyses, and for providing additional insight during interpretation of results.

Incorporating the results from assessments of indicators of affective states into statistical models could improve researchers’ ability to control for underlying affective influences on behaviour, thereby improving the accuracy of study results and interpretation (Loss et al., 2021). Additionally, recognising proxies of affective states as variables could meaningfully help explain more of the variation and consequently improve model fit, providing more precise statistical analyses that better account for individual variability (Loss et al., 2021). Integrating welfare assessments would therefore not only improve generalisability across study populations but could also ensure that differences in affect are properly accounted for in results interpretation. By applying integrating affect at the statistical level, researchers may also avoid overstandardization*, this is, over controlling experimental conditions (such as ensuring the same affective state in all individuals), which has also been discussed to contribute to poor reproducibility (Voelkl et al., 2020).

5.2 What data to collect

The increasing availability of non-disruptive* welfare assessment methods and advances in remote technologies like camera-traps and heat sensors should facilitate systematic integration of welfare assessments into wildlife studies (Browning, 2023; Browning et al., 2024). Captive animal welfare science studies have led to the identification of a number of behavioural (e.g., frequency of play behaviour; (Mintline et al., 2013), somatic (e.g., fur quality; (Bellanca et al., 2014), and cognitive (e.g. cognitive bias; (E. Bethell et al., 2012) indicators likely to reflect the affective states of animals. Advances in wild animal welfare science are also broadening the scope of available welfare indicators suitable for field studies, including evaluating stress using eye temperature monitoring (Bernat-Ponce et al., 2022), remotely assessing disease (Schilling et al., 2022) and other non-invasive methods (Beaulieu, 2024a, 2024b; Beaulieu & Masilkova, 2024; Bernat-Ponce et al., 2022; Browning et al., 2024; Jerem & Romero, 2023; Racciatti et al., 2022; Salas et al., 2024; Schilling et al., 2022; Watters et al., 2021).

To improve the reliability of welfare assessment methods and minimise bias, combining multiple validated indicators (triangulation) is recommended (Paul et al., 2022). This approach incorporates multiple independent indicators (e.g., physical, behavioural, somatic, and cognitive measures) to cross-validate findings and strengthen evaluations (Table 2). Indicators should also be validated for the species of interest to ensure appropriate application (Browning, 2023).

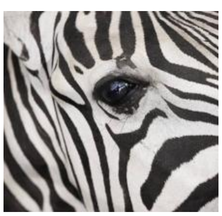
Using wild zebras as an example, Figure 1 illustrates some of the ways that affective status can be evaluated by combining various behavioural, physical, and cognitive observations recorded using non-invasive camera or video footage. In Table 2 we elaborate how such observations can be integrated with non-invasive collection of saliva (Baumann et al., 2025) and faecal samples. Such samples can be used for physiological analyses that provide information about immunological factors, disease load, stress levels and other physiological conditions affecting each animal. Table 2 provides an illustrative example of how investigators might use indicators of welfare to gain insights into affective motivations behind animal behaviours, using the example of observed differences in latency to move. Combining results from such observations and analyses of behavioural, cognitive, physical and physiological indicators provides a more informed understanding of affective states. By integrating assessment of affective states into studies, researchers are able to obtain greater insights into factors affecting patterns of behaviour.

Although integration of **these** types of data collection for evaluation of affective states may seem daunting, increased and systematic collection of welfare-relevant data itself will inherently lead to improvements in the methodologies, which will help streamline data collection and reduce challenges. Rapid technological advances such as AI and remote sensing will also help improve methods and efficiency, making it increasingly easier over time to systematically record and monitor welfare indicators as part of wildlife research.

It is also important to note that certain behavioural indicators, such as vigilance or allogrooming behaviours, are widely used in welfare assessments. However, the same behaviour should not be used as both the welfare indicator and the primary focus of the study.

Figure 1. Theoretical scenario of how to integrate welfare assessment, using wild zebras as an example. See Table 2 for indicators suggestive of behaviours, body condition, immunology, stress, etc. that give rise to potential interpretations.

Individual A	**Individual B**	**Individual C**	**Individual D**
Latency to move: High	**Latency to move: High**	**Latency to move: High**	**Latency to move: Low**
 Alert ears forward, head position lowered	 Aggressive facial expressions and vocalizations	 Ears forward, head position normal	 Pale gum colour = dehydration

 Signs of injury	 Scaring = signs of past predator attack	 Attention towards individual B	 Slow gait, Low head
 Increased attention and vigilance towards bushes (where predators are present)	 Increased attention and vigilance towards bushes (where predators are present)		
 Eyes alert and showing signs of agitation and discomfort.	 Pupils dilated	 Pupils not dilated	 Eyes clouded
 Head low, signs of agitation	 Allogrooming with Individual B	 Allogrooming with Individual B	

 Signs of fear	 Signs of fear		

Table 2. Results of the example scenario (Figure 1) showing the integration of welfare assessment in wild zebras, using physical and behavioural observations, along with physiological analyses from faecal samples and saliva swabs.

Welfare indicators	Latency	Ear position, Eye characteristics	Behaviours (posture, movement, micro-movements, aggressions, etc.)	Body Condition (injury presence, etc.)	Heart Rate / Temperature
Individual A	High	Ears, forward and flattened, pupils dilated, agitated, rapid eye movements,	Low body posture, reduced micro-movement, when moving gait is ‘abnormal’, steps are slowed	Signs of current injury (rear leg at awkward angle),	**High** eye temperature
Individual B	High	Pupil dilation, eyes alert, directionality neutral	Low posture, reduced micro-movement, when moving gait is ‘normal’, steps are not slowed, Aggressive behaviours (e.g., baring teeth), Associative behaviours towards individual C (e.g., nuzzling, allogrooming)	Signs of scarring (indicating previous interaction with predators), no signs of current injury	**High** eye temperature
Individual C	High	No pupil dilation, eyes alert	‘Normal’ posture, gait normal, Associative behaviours towards individual B (e.g., nuzzling, allogrooming)	No signs of scarring, injury, etc.	**No** raised eye temperature
Individual D	Low	Pupil dilation, sunken eyes, eyes clouded over	When moving gait is ‘abnormal’, steps are slowed but active (urgent), lethargic movements, head slung low,	No signs of scarring, injury, etc.	**High** eye temperature, sunken eyes, no sweating,

Welfare indicators (cont.)	Vigilance	Vocalizations	Attention Bias	Signs of dehydration	Immunology	Faecal glucocorticoids	Interpretations
Individual A	Low	Low whinnying	Focused on bushes	No signs of dehydration	Indications of immunological	Indications of stress	Latency to move is related

					response		to injury
Individual B	Increased alertness	Aggressive vocalizations	Towards bushes (where predators are present)	No signs of dehydration	No immunological indication of disease	Signs of increased negatively-valenced stress response.	Latency to move is related to increased fear response / predator avoidance
Individual C	No increased vigilance	No vocalizations	Towards individual B	No signs of dehydration	No immunological indication of disease	No signs of stress	Latency to move is related to association with Individual B.
Individual D	Increased vigilance	No vocalizations	Towards direction of travel (watering hole)	Signs of dehydration: dry gums (lack of saliva evident on imagery),	No immunological indication of disease	Increased stress response	Latency to move is related to individuals suffering from extensive dehydration and therefore has increased motivation to move to find water sources

5.4 Challenges for integrating welfare assessment

Evaluation of affective states uses proxies rather than direct measurement; thus, a persistent challenge is the absence of species-specific reliable and validated indicators of affective states. Moreover, because affective states are inherently subjective and thus inaccessible to direct measurement, it is not a true measure of animal experiences. In addition, no single indicator can offer complete accuracy; even the most robust and transferable measures inevitably require species-specific validation. This can complicate application of welfare assessment protocols for wildlife studies concerning species for which indicators have yet to be applied or validated (e.g.,(Harvey et al., 2023b).

Including affect as a variable substantially increases the complexity of statistical analyses. A larger sample size is likely to be required because adding affective state (which is inferred from collection of multiple welfare indicators) as a variable increases the number of parameters in the statistical models, which heightens the risk of overfitting. With more parameters, the model can begin to fit random noise within the data rather than genuine underlying patterns, reducing its validity and generalisability. When welfare is assessed with multiple indicators, especially when using triangulation approaches as suggested, the risk intensifies, as the model's complexity grows further. However, data with less variation can sometimes be achieved by collecting repeated measurements over time rather than increasing subject numbers.

In wild populations, additional challenges arise because numerous ecological, social, and individual-level factors fluctuate naturally over time. Variables such as environmental conditions, group composition, resource availability, and seasonal effects can introduce substantial variability that interacts with welfare indicators, making statistical interpretation more complicated. These inherent fluctuations need to be carefully considered when designing analyses and interpreting results, often necessitating more sophisticated modelling approaches or longitudinal study designs.

Relationships and redundancy between various welfare indicators, and other variables (e.g., sex, age, body condition) could also be a challenge or an efficiency. Nevertheless, the substantial explanatory value offered by these additional variables likely justifies the increased methodological effort.

Finally, integrating affect into animal research, whether through welfare assessment or by considering it within the interpretation of behavioural outcomes, will inevitably require additional time investment and close interdisciplinary collaboration.

6. Conclusion

In animal-based studies, the affective state of individual subjects is often regarded solely as an ethical consideration rather than an integral component of study design and statistical analysis. Not including an evaluation of individual affective states could leave substantial unnecessary unexplained variability, potentially skewing sampling, reducing generalisability, and undermining reproducibility, particularly in research conducted in wild populations. To contribute to enhancing the methodological rigour and replicability of animal studies both in the captive and wild environments, we recommend systematically incorporating welfare assessments into study protocols, and incorporating affective states as additional variables in statistical analyses.

Including assessment of welfare-related variables in studies would reveal key affective influences on animal responses to stimuli, and could substantially reduce unexplained variability, ensuring that findings are interpreted within a more transparent and replicable framework.

Integrating methods and lessons from animal welfare science and from good-welfare practices with rapidly advancing non-invasive, remote and continuous wildlife monitoring methodologies should help facilitate systematic inclusion of data collection for welfare assessments in wildlife studies.

Routine integration of welfare assessments into wildlife research will also help generate more relevant data for answering critical ecological questions about wildlife and inform our understanding of the challenges they face under anthropogenic and climate change.

7. References

Anderson, D. J. (2016). Circuit modules linking internal states and social behaviour in flies and mice. *Nature Reviews Neuroscience*, *17*(11), 692–704. https://doi.org/10.1038/nrn.2016.125

Anvari, F., & Lakens, D. (2019). *The Replicability Crisis and Public Trust in Psychological Science*. OSF. https://doi.org/10.31234/osf.io/vtmpc

Baciadonna, L., Bueno, J. C., Elliott, V., & Nawroth, C. (2024). Taking Welfare into Account in Comparative Cognition Research. *Comparative Cognition & Behavior Reviews*, *19*, 3–7.

Baker, M. (2016). 1,500 scientists lift the lid on reproducibility. *Nature*, *533*(7604), 452–454. https://doi.org/10.1038/533452a

Bartoš, F., Maier, M., Wagenmakers, E.-J., Nippold, F., Doucouliagos, H., Ioannidis, J. P. A., Otte, W. M., Sladekova, M., Deresssa, T. K., Bruns, S. B., Fanelli, D., & Stanley, T. D. (2024). Footprint of publication selection bias on meta-analyses in medicine, environmental sciences, psychology, and economics. *Research Synthesis Methods*, *15*(3), 500–511. https://doi.org/10.1002/jrsm.1703

Baumann, S., Euring, B., Harzer, M., Eibisch, M., Lindner, A., Vahlenkamp, T. W., & Heenemann, K. (2025). Development of a Novel, Non-Invasive Saliva Sampling Method for the Detection of Bovine Respiratory Viruses. *Veterinary Sciences*, *12*(7), 637. https://doi.org/10.3390/vetsci12070637

Beaulieu, M. (2024a). Capturing wild animal welfare: A physiological perspective. *Biological Reviews*, *99*(1), 1–22. https://doi.org/10.1111/brv.13009

Beaulieu, M. (2024b). Oxidative status: A general but overlooked indicator of welfare across animal species? *BioEssays*, *46*(8), 2300205. https://doi.org/10.1002/bies.202300205

Beaulieu, M., & Masilkova, M. (2024). Plugging biologging into animal welfare: An opportunity for advancing wild animal welfare science. *Methods in Ecology and Evolution*, *15*(12), 2172–2188. https://doi.org/10.1111/2041-210X.14441

Beausoleil, N. J., Mellor, D. J., Baker, L., Baker, S. E., Bellio, M., Clarke, A. S., Dale, A., Garlick, S., Jones, B., Harvey, A., Pitcher, B. J., Sherwen, S., Stockin, K. A., & Zito, S. (2018). "Feelings and Fitness" not "Feelings or Fitness" – The raison d'être of conservation welfare, which aligns conservation and animal welfare objectives. *Frontiers in Veterinary Science*, *5*. https://doi.org/10.3389/fvets.2018.00296

Begley, C. G., & Ellis, L. M. (2012). Raise standards for preclinical cancer research. *Nature*, *483*(7391), 531–533. https://doi.org/10.1038/483531a

Bellanca, R. U., Lee, G. H., Vogel, K., Ahrens, J., Kroeker, R., Thom, J. P., & Worlein, J. M. (2014). A simple alopecia scoring system for use in colony management of laboratory-housed primates. *Journal of Medical Primatology*, *43*(3), 153–161. https://doi.org/10.1111/jmp.12107

Beran, M. (2020). Editorial: The Value and Status of Replications in Animal Behavior and Cognition Research. *Animal Behavior and Cognition*, *7*(1), i–iii. https://doi.org/10.26451/abc.07.01.01.2020

Bernat-Ponce, E., MacLeod, R., & Nager, R. G. (2022). *Efficient non-invasive monitoring of an ecologically relevant physiological trait using thermal imaging of wild birds*. http://hdl.handle.net/11268/12386

Bethell, E., Holmes, A., MacLarnon, A., & Semple, S. (2012). Cognitive bias in a non-human primate: Husbandry procedures influence cognitive indicators of psychological well-being in captive rhesus macaques. *Animal Welfare*, *21*(2), 185–195. https://doi.org/10.7120/09627286.21.2.185

Bethell, E. J. (2015). A "how-to" guide for designing judgment bias studies to assess captive animal welfare. *Journal of Applied Animal Welfare Science*, *18*(sup1), S18–S42. https://doi.org/10.1080/10888705.2015.1075833

Boissy, A., Arnould, C., Chaillou, E., Désiré, L., Duvaux-Ponter, C., Greiveldinger, L., Leterrier, C., Richard, S., Roussel, S., Saint-Dizier, H., Meunier-Salaün, M. C., Valance, D., & Veissier, I. (2007). Emotions and cognition: A new approach to animal welfare. *Animal Welfare*, *16*(S1), 37–

43. https://doi.org/10.1017/S0962728600031717

Brecht, K. F., Legg, E. W., Nawroth, C., Fraser, H., & Ostojic, L. (2021). The Status and Value of Replications in Animal Behavior Science. *Animal Behavior and Cognition*, *8*(2), 97–106. https://doi.org/10.26451/abc.08.02.01.2021

Brill, S. A., Guerrero-Martin, S. M., & Metcalf Pate, K. A. (2019). The Symbiotic Relationship Between Scientific Quality and Animal Research Ethics. *ILAR Journal*, *60*(3), 334–340. https://doi.org/10.1093/ilar/ilab023

Browning, H. (2022). Assessing measures of animal welfare. *Biology & Philosophy*, *37*(4), 36. https://doi.org/10.1007/s10539-022-09862-1

Browning, H. (2023). Validating Indicators of Subjective Animal Welfare. *Philosophy of Science*, *90*(5), 1255–1264. https://doi.org/10.1017/psa.2023.10

Browning, H., Beaulieu, M., Sharo, A., Barrett, M., Flint, B., Castellano Bueno, J., Clay, A., Veit, W., & Elliott, V. (2024). *Assessing the suitability of welfare indicators for wild animals*. https://osf.io/k3scb

Bruns, H., Dessart, F. J., Krawczyk, M., Lewandowsky, S., Pantazi, M., Pennycook, G., Schmid, P., & Smillie, L. (2024). Investigating the role of source and source trust in prebunks and debunks of misinformation in online experiments across four EU countries. *Scientific Reports*, *14*, 20723. https://doi.org/10.1038/s41598-024-71599-6

Budaev, S., Jørgensen, C., Mangel, M., Eliassen, S., & Giske, J. (2019). Decision-Making From the Animal Perspective: Bridging Ecology and Subjective Cognition. *Frontiers in Ecology and Evolution*, *7*. https://doi.org/10.3389/fevo.2019.00164

Bushby, E. V., Friel, M., Goold, C., Gray, H., Smith, L., & Collins, L. M. (2018). Factors Influencing Individual Variation in Farm Animal Cognition and How to Account for These Statistically. *Frontiers in Veterinary Science*, *5*, 193. https://doi.org/10.3389/fvets.2018.00193

Carranza-Pinedo, V., Krohs, U., & Richter, S. H. (2025). Towards a scientific definition of animal emotions: Integrating innate, appraisal, and network mechanisms. *Neuroscience &*

Biobehavioral Reviews, *172*, 106127. https://doi.org/10.1016/j.neubiorev.2025.106127

Cassidy, L. C., Hannibal, D. L., Semple, S., & McCowan, B. (2020). Improved behavioral indices of welfare in continuous compared to intermittent pair-housing in adult female rhesus macaques (Macaca mulatta). *American Journal of Primatology*, *82*(10), e23189. https://doi.org/10.1002/ajp.23189

Chambers, C. D., & Tzavella, L. (2022). The past, present and future of Registered Reports. *Nature Human Behaviour*, *6*(1), 29–42. https://doi.org/10.1038/s41562-021-01193-7

Cheleuitte-Nieves, C., & Lipman, N. S. (2019). Improving Replicability, Reproducibility, And Reliability In Preclinical Research: A Shared Responsibility. *ILAR Journal*, *60*(2), 113–119. https://doi.org/10.1093/ilar/ilaa009

Cook, C. N., Freeman, A. R., Liao, J. C., & Mangiamele, L. A. (2021). The Philosophy of Outliers: Reintegrating Rare Events Into Biological Science. *Integrative and Comparative Biology*, *61*(6), 2191–2198. https://doi.org/10.1093/icb/icab166

Crump, A., Arnott, G., & Bethell, E. J. (2018). Affect-Driven Attention Biases as Animal Welfare Indicators: Review and Methods. *Animals*, *8*(8), Article 8. https://doi.org/10.3390/ani8080136

Das, R. E. G. (2002). Role of Ancillary Variables in the Design, Analysis, and Interpretation of Animal Experiments. *ILAR Journal*, *43*(4), 214–222. https://doi.org/10.1093/ilar.43.4.214

Dawkins, M. S. (2004). Using behaviour to assess animal welfare. *Animal Welfare*, *13*(S1), S3–S7. https://doi.org/10.1017/S0962728600014317

Descovich, K. A., Richmond, S. E., Leach, M. C., Buchanan-Smith, H. M., Flecknell, P., Farningham, D. A. H., Witham, C., Gates, M. C., & Vick, S.-J. (2019). Opportunities for refinement in neuroscience: Indicators of wellness and post-operative pain in laboratory macaques. *ALTEX*, *36*(4), 535–554. https://doi.org/10.14573/altex.1811061

Directive 2010/63/EU of the European Parliament and of the Council of 22 September 2010 on the protection of animals used for scientific purposesText with EEA relevance. (n.d.).

Eberhart-Hertel, L., Williams, E. M., McGilvary-Howard, A., Howard, T., Habraken, T., O'Donnell, C. F. J., Küpper, C., & Kempenaers, B. (2025). Evaluating the effects of tracking devices on survival, breeding success, behavior, and condition of a small, partially migratory shorebird. *Journal of Avian Biology*, *2025*(4), e03490. https://doi.org/10.1002/jav.03490

Ekkekakis, P. (2013). *The Measurement of Affect, Mood, and Emotion: A Guide for Health-Behavioral Research*. Cambridge University Press. https://doi.org/10.1017/CBO9780511820724

Eldegard, K., Furnes, M. W., Grainger, M., Moe, B., Sandercock, B. K., Sonerud, G. A., Ytrehus, B., Rueness, E. K., Sayyari, A., Kirkendall, L. R., Granquist, E. G., & Kausrud, K. (2024). Effects of capture, marking, and tracking on the welfare of wild birds. *227*. https://nordopen.nord.no/nord-xmlui/handle/11250/3178904

Fidler, F., Chee, Y. E., Wintle, B. C., Burgman, M. A., McCarthy, M. A., & Gordon, A. (2017). Metaresearch for Evaluating Reproducibility in Ecology and Evolution. *BioScience*, *67*(3), 282–289. https://doi.org/10.1093/biosci/biw159

Filazzola, A., & Cahill Jr, J. F. (2021). Replication in field ecology: Identifying challenges and proposing solutions. *Methods in Ecology and Evolution*, *12*(10), 1780–1792. https://doi.org/10.1111/2041-210X.13657

Florczak, K. L. (2022). Best Available Evidence or Truth for the Moment: Bias in Research. *Nursing Science Quarterly*, *35*(1), 20–24. https://doi.org/10.1177/08943184211051350

Franco, N. H., Magalhães-Sant'Ana, M., & Olsson, I. A. S. (2014). Welfare and quantity of life. In *Dilemmas in animal welfare* (pp. 46–66). https://doi.org/10.1079/9781780642161.0046

Freedman, L. P., Cockburn, I. M., & Simcoe, T. S. (2015). The Economics of Reproducibility in Preclinical Research. *PLOS Biology*, *13*(6), e1002165. https://doi.org/10.1371/journal.pbio.1002165

Geldart, E. A., Howes, L.-A., Wheeler, H., & Mackenzie, S. A. (n.d.). *A review of Impacts of Tracking Devices on Birds*.

Gregol, R. K., Souza, D. N. de, & Verrastro, L. (2024). *Behavioral description of the endangered lizard Liolaemus arambarensis in its natural Restinga environment through continuous filming |*

REVISTA DELOS. https://ojs.revistadelos.com/ojs/index.php/delos/article/view/3147

Grimm, D. (2018). *Are happy lab animals better for science?* [Data set]. https://doi.org/10.1126/science.aat2326

Guerrero-Martin, S. M., Rubin, L. H., McGee, K. M., Shirk, E. N., Queen, S. E., Li, M., Bullock, B., Carlson, B. W., Adams, R. J., Gama, L., Graham, D. R., Zink, C., Clements, J. E., Mankowski, J. L., & Metcalf Pate, K. A. (2021). Psychosocial Stress Alters the Immune Response and Results in Higher Viral Load During Acute Simian Immunodeficiency Virus Infection in a Pigtailed Macaque Model of Human Immunodeficiency Virus. *The Journal of Infectious Diseases*, *224*(12), 2113–2121. https://doi.org/10.1093/infdis/jiab252

Guttinger, S. (2020). The limits of replicability. *European Journal for Philosophy of Science*, *10*(2), 10. https://doi.org/10.1007/s13194-019-0269-1

Harvey, A. M., Beausoleil, N. J., Ramp, D., & Mellor, D. J. (2023a). Mental experiences in wild animals: Scientifically validating measurable welfare indicators in free-roaming horses. *Animals*, *13*(9), 1507. https://doi.org/10.3390/ani13091507

Harvey, A. M., Beausoleil, N. J., Ramp, D., & Mellor, D. J. (2023b). Mental Experiences in Wild Animals: Scientifically Validating Measurable Welfare Indicators in Free-Roaming Horses. *Animals*, *13*(9), 1507. https://doi.org/10.3390/ani13091507

Hawkins, P., Morton, D. B., Burman, O., Dennison, N., Honess, P., Jennings, M., Lane, S., Middleton, V., Roughan, J. V., Wells, S., & Westwood, K. (2011). A guide to defining and implementing protocols for the welfare assessment of laboratory animals: Eleventh report of the BVAAWF/FRAME/RSPCA/UFAW Joint Working Group on Refinement*. *Laboratory Animals*, *45*(1), 1–13. https://doi.org/10.1258/la.2010.010031

Hecht, L. (2021). The importance of considering age when quantifying wild animals’ welfare. *Biological Reviews*, *96*(6), 2602–2616. https://doi.org/10.1111/brv.12769

Jerem, P., & Romero, L. M. (2023). It’s cool to be stressed: Body surface temperatures track sympathetic nervous system activation during acute stress. *The Journal of Experimental Biology*, *226*(20), jeb246552. https://doi.org/10.1242/jeb.246552

Jilka, R. L. (2016). The Road to Reproducibility in Animal Research. *Journal of Bone and Mineral Research*, *31*(7), 1317–1319. https://doi.org/10.1002/jbmr.2881

Kennedy, A., Asahina, K., Hoopfer, E., Inagaki, H., Jung, Y., Lee, H., Remedios, R., & Anderson, D. J. (2014). Internal States and Behavioral Decision-Making: Toward an Integration of Emotion and Cognition. *Cold Spring Harbor Symposia on Quantitative Biology*, *79*, 199–210. https://doi.org/10.1101/sqb.2014.79.024984

Kervinen, M., Lebigre, C., & Soulsbury, C. D. (2016). Simultaneous age-dependent and age-independent sexual selection in the lekking black grouse (Lyrurus tetrix). *Journal of Animal Ecology*, *85*(3), 715–725. https://doi.org/10.1111/1365-2656.12496

Kinnally, E. L., Martinez, S. J., Chun, K., Capitanio, J. P., & Ceniceros, L. C. (2019). Early Social Stress Promotes Inflammation and Disease Risk in Rhesus Monkeys. *Scientific Reports*, *9*(1), 7609. https://doi.org/10.1038/s41598-019-43750-1

Kitajima, K., Oishi, K., Miwa, M., Anzai, H., Setoguchi, A., Yasunaka, Y., Himeno, Y., Kumagai, H., & Hirooka, H. (2021). Effects of Heat Stress on Heart Rate Variability in Free-Moving Sheep and Goats Assessed With Correction for Physical Activity. *Frontiers in Veterinary Science*, *8*, 658763. https://doi.org/10.3389/fvets.2021.658763

LaRochelle, L., Griffin, L. P., Brownscombe, J. W., Elvidge, C. K., Hasler, C. T., Robichaud, J. A., Madden, J. C., Landsman, S. J., Raoult, V., Danylchuk, A. J., & Cooke, S. J. (2025). On the behavior of Fish Released Following Fisheries Capture: Methods, Endpoints, and Consequences. *Reviews in Fisheries Science & Aquaculture*. https://www.tandfonline.com/doi/abs/10.1080/23308249.2025.2518167

Lee, C., Verbeek, E., Doyle, R., & Bateson, M. (2016). Attention bias to threat indicates anxiety differences in sheep. *Biology Letters*, *12*(6), 20150977. https://doi.org/10.1098/rsbl.2015.0977

Lench, H. C., Flores, S. A., & Bench, S. W. (2011). Discrete emotions predict changes in cognition, judgment, experience, behavior, and physiology: A meta-analysis of experimental emotion elicitations. *Psychological Bulletin*, *137*(5), 834–855. https://doi.org/10.1037/a0024244

Lerner, J. S., Li, Y., Valdesolo, P., & Kassam, K. S. (2015). Emotion and Decision Making. *Annual Review of Psychology*, *66*(1), 799–823. https://doi.org/10.1146/annurev-psych-010213-115043

Li, Y., Zhang, B., Yang, Y., Su, P., Samsom, J. N., Wong, A. H. C., & Liu, F. (2024). Sex and Age Differences in Glucocorticoid Signaling After an Aversive Experience in Mice. *Cells*, *13*(24), 2041. https://doi.org/10.3390/cells13242041

Loss, C. M., Melleu, F. F., Domingues, K., Lino-de-Oliveira, C., & Viola, G. G. (2021). Combining Animal Welfare With Experimental Rigor to Improve Reproducibility in Behavioral Neuroscience. *Frontiers in Behavioral Neuroscience*, *15*. https://doi.org/10.3389/fnbeh.2021.763428

MacArthur Clark, J. A., & Sun, D. (2020). Guidelines for the ethical review of laboratory animal welfare People's Republic of China National Standard GB/T 35892-2018 [Issued 6 February 2018 Effective from 1 September 2018]. *Animal Models and Experimental Medicine*, *3*(1), 103–113. https://doi.org/10.1002/ame2.12111

MacLellan, A., & Lalu, M. M. (2024). Testing the limits: Can enhanced external validity promote translation? *eBioMedicine*, *110*, 105473. https://doi.org/10.1016/j.ebiom.2024.105473

MacLellan, A., Nazal, B., Young, L., & Mason, G. (2022). Waking inactivity as a welfare indicator in laboratory mice: Investigating postures, facial expressions and depression-like states. *Royal Society Open Science*, *9*(11), 221083. https://doi.org/10.1098/rsos.221083

Marchant-Forde, J. N. (2015). The Science of Animal Behavior and Welfare: Challenges, Opportunities, and Global Perspective. *Frontiers in Veterinary Science*, *2*. https://doi.org/10.3389/fvets.2015.00016

Mason, G., & Bateson, M. (2017). Motivation and the organization of behaviour. In *The ethology of domestic animals: An introductory text* (pp. 38–61). https://doi.org/10.1079/9781786391650.0038

Mellor, D. (2015). Enhancing animal welfare by creating opportunities for positive affective engagement. *New Zealand Veterinary Journal*, *63*(1), 3–8. https://doi.org/10.1080/00480169.2014.926799

Mellor, D. J. (2016). Updating animal welfare thinking: Moving beyond the “five freedoms” towards “A life worth living”. *Animals*, *6*(3). https://doi.org/10.3390/ani6030021

Mendl, M., Burman, O. H. P., Parker, R. M. A., & Paul, E. S. (2009). Cognitive bias as an indicator of animal emotion and welfare: Emerging evidence and underlying mechanisms. Applied Animal Behaviour Science, 118(3–4), 161–181.

Mintline, E. M., Stewart, M., Rogers, A. R., Cox, N. R., Verkerk, G. A., Stookey, J. M., Webster, J. R., & Tucker, C. B. (2013). Play behavior as an indicator of animal welfare: Disbudding in dairy calves. *Applied Animal Behaviour Science*, *144*(1), 22–30. https://doi.org/10.1016/j.applanim.2012.12.008

Monk, J. E., Lee, C., Dickson, E., & Campbell, D. L. M. (2020). Attention Bias Test Measures Negative But Not Positive Affect in Sheep: A Replication Study. *Animals: An Open Access Journal from MDPI*, *10*(8), 1314. https://doi.org/10.3390/ani10081314

Montalcini, Camille M., Charles C. Driver, and Michael T. Mendl. “Relevance of State-Behaviour Feedbacks for Animal Welfare.” *Biological Reviews of the Cambridge Philosophical Society* 100, no. 4 (2025): 1615–34. https://doi.org/10.1111/brv.70016.

Mundinger, C., Schulz, N. K. E., Singh, P., Janz, S., Schurig, M., Seidemann, J., Kurtz, J., Müller, C., Schielzeth, H., Kortzfleisch, V. T. von, & Richter, S. H. (2025). Testing the reproducibility of ecological studies on insect behavior in a multi-laboratory setting identifies opportunities for improving experimental rigor. *PLOS Biology*, *23*(4), e3003019. https://doi.org/10.1371/journal.pbio.3003019

Nakagawa, S., & Parker, T. H. (2015). Replicating research in ecology and evolution: Feasibility, incentives, and the cost-benefit conundrum. *BMC Biology*, *13*(1), 88. https://doi.org/10.1186/s12915-015-0196-3

Nawroth, C., & Gygax, L. (2020). *The legislative, ethical, and conceptual importance of replicability in farm animal welfare science.* agriRxiv. https://doi.org/10.31220/osf.io/ahspd

Niella, Y., Simes, B., Fox, A., Wright, A., Waller, M., Riley, M., Meyer, L., Drew, M., Pederson, H., & Huveneers, C. (2023). Short-term response of research activities on white shark behaviour. *Wildlife Research*, *50*(4), 260–271. https://doi.org/10.1071/WR22004

Niśkiewicz, M., Szymański, P., Zampa, L., Budka, M., & Osiejuk, T. (2024). Response of the emerald-spotted wood-dove to the song of conspecific males and sympatric congeners. *Ethology*, *130*, e13498. https://doi.org/10.1111/eth.13498

Panksepp, J. (2011). Toward a cross-species neuroscientific understanding of the affective mind: Do animals have emotional feelings? *American Journal of Primatology*, *73*(6), 545–561. https://doi.org/10.1002/ajp.20929

Papini, M. R., Penagos-Corzo, J. C., & Pérez-Acosta, A. M. (2019). Avian Emotions: Comparative Perspectives on Fear and Frustration. *Frontiers in Psychology*, *9*. https://doi.org/10.3389/fpsyg.2018.02707

Parker, T. H., Forstmeier, W., Koricheva, J., Fidler, F., Hadfield, J. D., Chee, Y. E., Kelly, C. D., Gurevitch, J., & Nakagawa, S. (2016). Transparency in Ecology and Evolution: Real Problems, Real Solutions. *Trends in Ecology & Evolution*, *31*(9), 711–719. https://doi.org/10.1016/j.tree.2016.07.002

Paul, E. S., Browne, W., Mendl, M. T., Caplen, G., Trevarthen, A., Held, S., & Nicol, C. J. (2022). Assessing animal welfare: A triangulation of preference, judgement bias and other candidate welfare indicators. *Animal Behaviour*, *186*, 151–177. https://doi.org/10.1016/j.anbehav.2022.02.003

Paul, E. S., Harding, E. J., & Mendl, M. (2005). Measuring emotional processes in animals: The utility of a cognitive approach. *Neuroscience and Biobehavioral Reviews*, *29*(3), 469–491. https://doi.org/10.1016/j.neubiorev.2005.01.002

Percie Du Sert, N., Hurst, V., Ahluwalia, A., Alam, S., Avey, M. T., Baker, M., Browne, W. J., Clark, A., Cuthill, I. C., Dirnagl, U., Emerson, M., Garner, P., Holgate, S. T., Howells, D. W., Karp, N. A., Lazic, S. E., Lidster, K., MacCallum, C. J., Macleod, M., … Wuerbel, H. (2020). The ARRIVE guidelines 2.0: Updated guidelines for reporting animal researchThe ARRIVE guidelines 2.0: updated guidelines for reporting animal research. *BMJ Open Science*, *44*(11). https://doi.org/10.1136/bmjos-2020-100115

Percie du Sert, N., Hurst, V., Ahluwalia, A., Alam, S., Avey, M. T., Baker, M., Browne, W. J., Clark, A., Cuthill, I. C., Dirnagl, U., Emerson, M., Garner, P., Holgate, S. T., Howells, D. W., Karp, N. A., Lidster, K., MacCallum, C. J., Macleod, M., Petersen, O., … Würbel, H. (2019). The ARRIVE guidelines 2019: Updated guidelines for reporting animal research. *bioRxiv*. https://doi.org/10.1101/703181

Poole, T. (1997). Happy animals make good science. *Laboratory Animals*, *31*(2), 116–124. https://doi.org/10.1258/002367797780600198

Powers, S. M., & Hampton, S. E. (2019). Open science, reproducibility, and transparency in ecology. *Ecological Applications*, *29*(1), e01822. https://doi.org/10.1002/eap.1822

Prescott, M. J., & Lidster, K. (2017). Improving quality of science through better animal welfare: The NC3Rs strategy. *Lab Animal*, *46*(4), 152–156. https://doi.org/10.1038/laban.1217

Racciatti, D. S., Feld, A., Rial, L. A., Blanco, C., & Tallo-Parra, O. (2022). Ackonc-AWA: A multi-species animal welfare assessment protocol for wild animals under human care to overcome the use of generic welfare checklists. *Frontiers in Veterinary Science*, *9*. https://doi.org/10.3389/fvets.2022.1033821

Richter, S. H. (2017). Systematic heterogenization for better reproducibility in animal experimentation. *Lab Animal*, *46*(9), 343–349. https://doi.org/10.1038/laban.1330

Richter, S. H., Garner, J. P., & Würbel, H. (2009). Environmental standardization: Cure or cause of poor reproducibility in animal experiments? *Nature Methods*, *6*(4), 257–261. https://doi.org/10.1038/nmeth.1312

Russell, W. M. S., & Burch, R. L. (1959). *The Principles of Humane Experimental Technique.*

Methuen.

Salas, M., Tallo-Parra, O., & Manteca, X. (2024). Evidence-based zoo animal welfare assessment: Putting science into practice. *Journal of Zoo and Aquarium Research*, *12*(4), 205–211. https://doi.org/10.19227/jzar.v12i4.810

Schad, L., van de Waal, E., & Fischer, J. (2025). Anti-Snake Behavior and Snake Discrimination in Vervet Monkeys. *Ethology*, *131*(3), e13541. https://doi.org/10.1111/eth.13541

Schilling, A.-K., Mazzamuto, M. V., & Romeo, C. (2022). A Review of Non-Invasive Sampling in Wildlife Disease and Health Research: What's New? *Animals*, *12*(13), 1719. https://doi.org/10.3390/ani12131719

Schloss, P. D. (2018). Identifying and Overcoming Threats to Reproducibility, Replicability, Robustness, and Generalizability in Microbiome Research. *mBio*, *9*(3), e00525-18. https://doi.org/10.1128/mBio.00525-18

Shaw, R. C., Greggor, A. L., & Plotnik, J. M. (2021). The Challenges of Replicating Research on Endangered Species. *Animal Behavior and Cognition*, *8*(2), 240–246. https://doi.org/10.26451/abc.08.02.10.2021

Sleegers, Willem W. A., David Moss, William H. B. McAuliffe, David Reinstein, and Daniela R. Waldhorn. "Measuring Wild Animal Welfare Attitudes: The Attitudes Towards Wild Animal Welfare Scale." Psychology of Human-Animal Intergroup Relations 4 (October 2025): 1–26. https://doi.org/10.5964/phair.17951.

Smith, A. J., Clutton, R. E., Lilley, E., Hansen, K. E. A., & Brattelid, T. (2018). PREPARE: Guidelines for planning animal research and testing. *Laboratory Animals*, *52*(2), 135–141. https://doi.org/10.1177/0023677217724823

Smith, G. D., & Ebrahim, S. (2002). Data dredging, bias, or confounding. *BMJ : British Medical Journal*, *325*(7378), 1437–1438. https://doi.org/10.1136/bmj.325.7378.1437

Sockman, K. W., & Beaulieu, M. (2023). Paradoxical associations between fitness components and behavioural phenotype in a wild bird. *Animal Behaviour*, *198*, 47–57.

https://doi.org/10.1016/j.anbehav.2023.01.015

Sosa, S. (2016). The Influence of Gender, Age, Matriline and Hierarchical Rank on Individual Social Position, Role and Interactional Patterns in Macaca sylvanus at 'La Forêt des Singes': A Multilevel Social Network Approach. *Frontiers in Psychology*, *7*. https://doi.org/10.3389/fpsyg.2016.00529

Takahashi, L. K. (1986). Postweaning environmental and social factors influencing the onset and expression of agonistic behavior in Norway rats. *Behavioural Processes*, *12*(3), 237–260. https://doi.org/10.1016/0376-6357(86)90039-2

Truelove, M. A., Martin, J. E., Langford, F. M., & Leach, M. C. (2020). The identification of effective welfare indicators for laboratory-housed macaques using a Delphi consultation process. *Scientific Reports*, *10*(1), 20402. https://doi.org/10.1038/s41598-020-77437-9

Veissier, I., Aubert, A., & Boissy, A. (2012). Animal welfare: A result of animal background and perception of its environment. *Animal Frontiers*, *2*(3), 7–15. https://doi.org/10.2527/af.2012-0043

Voelkl, B., Altman, N. S., Forsman, A., Forstmeier, W., Gurevitch, J., Jaric, I., Karp, N. A., Kas, M. J., Schielzeth, H., Van de Casteele, T., & Würbel, H. (2020). Reproducibility of animal research in light of biological variation. *Nature Reviews Neuroscience*, *21*(7), 384–393. https://doi.org/10.1038/s41583-020-0313-3

Voelkl, B., Vogt, L., Sena, E. S., & Würbel, H. (2018). Reproducibility of preclinical animal research improves with heterogeneity of study samples. *PLOS Biology*, *16*(2), e2003693. https://doi.org/10.1371/journal.pbio.2003693

Volsche, S., Root-Gutteridge, H., Korzeniowska, A. T., & Horowitz, A. (2023). Centring individual animals to improve research and citation practices. *Biological Reviews*, *98*(2), 421–433. https://doi.org/10.1111/brv.12912

von Borell, E., Langbein, J., Després, G., Hansen, S., Leterrier, C., Marchant, J., Marchant-Forde, R., Minero, M., Mohr, E., Prunier, A., Valance, D., & Veissier, I. (2007). Heart rate variability as a measure of autonomic regulation of cardiac activity for assessing stress and welfare in farm animals—A review. *Physiology & Behavior*, *92*(3), 293–316. https://doi.org/10.1016/j.physbeh.2007.01.007

Watters, J. V., Krebs, B. L., & Eschmann, C. L. (2021). Assessing Animal Welfare with Behavior: Onward with Caution. *Journal of Zoological and Botanical Gardens*, *2*(1), Article 1. https://doi.org/10.3390/jzbg2010006

Webster, M. M., & Rutz, C. (2020). How STRANGE are your study animals? *Nature*, *582*(7812), 337–340. https://doi.org/10.1038/d41586-020-01751-5

Weiß, B. M., & Foerster, K. (2013). Age and sex affect quantitative genetic parameters for dominance rank and aggression in free-living greylag geese. *Journal of Evolutionary Biology*, *26*(2), 299–310. https://doi.org/10.1111/jeb.12042

Wicherts, Jelte M. "The Weak Spots in Contemporary Science (and How to Fix Them)." *Animals* 7, no. 12 (2017). https://doi.org/10.3390/ani7120090.

Williams, H. J., Taylor, L. A., Benhamou, S., Bijleveld, A. I., Clay, T. A., de Grissac, S., Demšar, U., English, H. M., Franconi, N., Gómez-Laich, A., Griffiths, R. C., Kay, W. P., Morales, J. M., Potts, J. R., Rogerson, K. F., Rutz, C., Spelt, A., Trevail, A. M., Wilson, R. P., & Börger, L. (2020). Optimizing the use of biologgers for movement ecology research. *Journal of Animal Ecology*, *89*(1), 186–206. https://doi.org/10.1111/1365-2656.13094

Yeates, J. W. (2024). Animal behaviour and welfare research: A One Health perspective. *Research Ethics*, *20*(3), 411–432. https://doi.org/10.1177/17470161241236941

Zemanova, M. A. (2021). New online resource on the 3Rs principles of animal research for wildlife biologists, ecologists, and conservation managers. *Conservation*, *1*(2), 106–112. https://doi.org/10.3390/conservation1020009