Applied Animal Behaviour Science 264 (2023) 105965

Contents lists available at ScienceDirect

Applied Animal Behaviour Science

journal homepage: www.elsevier.com/locate/applanim

Doing better for understudied species: Evaluation and improvement of a species-general animal welfare assessment tool for zoos

Shannon L. O'Brien *, Katherine A. Cronin

Lincoln Park Zoo, 2001 N. Clark Street, Chicago, IL 60614, USA

ARTICLE INFO

Keywords:
Animal welfare
Welfare assessment
Zoo(s)
Aquarium(s)
Taxonomic differences

ABSTRACT

Accredited zoos and aquariums are required to have a welfare assessment process in place for the animals in their care. However, welfare assessment theoretical frameworks and methodologies vary widely. Previous studies have shown that taxonomic biases in research are found broadly within the field of animal behavior and specifically within applied animal welfare science, which could translate to taxonomic biases in applied welfare assessment tools. After creating and then using a species-general welfare assessment at Lincoln Park Zoo for 2.5 years, we sought to examine whether there were systematic differences in welfare scores for different taxa. There were 939 Welfare Discussion Tools completed for 174 species between July 2019 and – December 2021, including mammals, birds, reptiles, amphibians, fish, and invertebrates. We found that, broadly, no taxonomic group scored higher than mammals on overall welfare scores, input (resource-based) scores, or output (animal-based) scores. Raters also provided fewer incomplete responses (in the form of "not applicable" or "I don't know" judgments) for mammals than for other taxonomic groups. Integrating these findings with input from welfare scientists, animal caretakers, behavioral husbandry experts and veterinary staff, we modified our original welfare assessment tool to broaden the taxonomic relevance of the tool. Copies of the original tool and the updated/ modified tool are included to serve as a model for institutions developing or revising their own welfare assessment tools.

1. Introduction

In recent decades, animal welfare science has transformed from a discipline interested in limiting animal suffering to one that strives to provide animals with environments and opportunities where they can thrive (Dawkins, 2021; Veasey, 2017). This has led to some significant advances within many zoos and aquariums, namely, the broad implementation of processes to methodically assess the welfare of animals in human care (Jones et al., 2022). Accrediting bodies such as the Association of Zoos and Aquariums, European Association of Zoos and Aquaria, and Zoo and Aquarium Association Australasia have made welfare assessments part of their required procedures and documentation when zoos or aquariums apply for accreditation (e.g., Association of Zoos and Aquariums, 2023; Zoo and Aquarium Association Australasia, 2023). This requirement necessitates that zoos and aquariums seeking accreditation have a welfare assessment process in place.

However, welfare assessment theoretical frameworks and methodologies vary widely (Jones et al., 2022) and therefore approaches for assessing animal welfare differ greatly between and within associations. For example, Jones et al. (2022) review many theoretical frameworks and welfare assessment approaches including the Five Freedoms model (Brambell, 1965), the Five Domains model (Mellor, 2017), the European Welfare Quality assessment protocol (Temple et al., 2012), the Universal Animal Welfare Framework (Kagan et al., 2015), the Opportunities to Thrive Program (Greggor et al., 2018), and the Animal Welfare Assessment Grid (Honess and Wolfensohn, 2010). Each of these frameworks and approaches differs in the metrics used for assessing animal welfare as well as the taxonomic breadth with which the assessment approach has been applied (Jones et al., 2022). In addition to the different theoretical frameworks, welfare assessment approaches may vary with regard to how they are applied (e.g., whether they are animal- or habitat-based, individual- or group-based, and input- or output-based) and how they are performed (e.g., who performs the assessment, how many people perform the assessment, and how often assessments are performed).

One common way in which welfare assessment approaches differ is with regard to whether the welfare assessment is species-general or species-specific. Species-general approaches are the most common

* Corresponding author.
E-mail address: sobrien@lpzoo.org (S.L. O'Brien).

https://doi.org/10.1016/j.applanim.2023.105965
Received 25 April 2023; Accepted 18 May 2023
Available online 19 May 2023

(Jones et al., 2022), although some species-specific tools are available in the published literature (e.g., giant pacific octopus: Holst and Miller-Morgan, 2021; waterfowl: Rose and O'Brien, 2020; pygmy blue tongued skink: Benn et al., 2019; bottle-nosed dolphin: Clegg et al., 2015), with most focusing on agricultural animals (Hill and Broom, 2009). The common use of species-general tools likely stems from the increased feasibility, as zoos and aquariums often house hundreds of different species that need to be assessed regularly to meet accreditation requirements.

Minimal guidance has been provided regarding which welfare assessment approach should be used and the frequency with which welfare assessments should take place for accreditation purposes. As zoos and aquariums approach welfare assessments differently, it is valuable to evaluate and share welfare assessment approaches to ensure they provide information that is both as accurate and as actionable as possible. Beyond this, we argue that specific attention to potential taxonomic biases is warranted, especially when using a species-general assessment approach that is broad and has less focus on species-specific welfare indicators.

Taxonomic bias in welfare assessment tools may emerge from a disproportionate amount of information available regarding the welfare of mammals compared to non-mammalian species (Binding et al., 2020; Jones et al., 2022). Historically, research in animal behavior has been biased toward vertebrates. A systematic review examined the taxa represented in publications between the years 1953 – 2015 from the journal *Animal Behaviour* and found that research on birds and mammals comprised over half of all publications (Rosenthal et al., 2017). As the authors note, this finding is particularly significant when one accounts for the fact that birds and mammals comprise less than 1.5% of all animal species. Similarly, a recent study that focused specifically on academic journals commonly publishing zoo and aquarium-based animal welfare research found that mammals comprised 75% of research between the years 2008 – 2017 (Binding et al., 2020). More broadly, the authors reported that 82% of studies were focused on vertebrates and there was a notable bias toward research on great apes. Collectively, these findings indicate that taxonomic biases in research are found broadly within the field of animal behavior and specifically within applied animal welfare science.

After using the species-general Welfare Discussion Tool at Lincoln Park Zoo for 2.5 years and assessing 174 different species, we sought to examine whether there were systematic differences in welfare scores for different taxa. Such differences could reflect a taxonomic bias in the development of the tool and/or a systematic difference in the welfare of different taxonomic groups in our care. Concurrently, we sought to gain general insights from our welfare assessment data to guide improvements to our welfare discussion tool. We share our welfare discussion tool, and importantly, our evaluation of our welfare discussion tool, to serve as a model for other institutions that are developing or revising their own welfare assessments.

2. Material and method

2.1. Welfare discussion tool

The Lincoln Park Zoo Welfare Discussion Tool (hereafter WDT) was created at Lincoln Park Zoo and informed by published species-general tools (e.g., Sherwen et al., 2018; Kagan et al., 2015). The WDT was developed with both input (resource-based) and output (animal-based) measures, acknowledging that the former are more readily measurable but the latter more informative about welfare states (Wolfensohn et al., 2018; Salas and Manteca 2016; Whitham and Wielebnowski, 2013).

The WDT allowed for quantitative scoring of agreement with statements/items, where more agreement (i.e., higher total welfare score) likely indicated better welfare, and acknowledged the lack of available objective data in many cases in two ways. First, the WDT allowed raters to respond "I don't know" or "not applicable" rather than force responses. Second, all raters participated in an in-person or virtual training meeting during which the philosophy behind the tool development and the mechanics of using the tool were discussed. During training, we communicated repeatedly that the tool was a "Welfare *Discussion* Tool" to highlight that objective measurements of welfare were potentially unrealistic for some items and species. We emphasized that an important step in the welfare assessment process was the discussion that occurred between three raters following completion of the tool.

The WDT consisted of 41 items (statements) categorized as either inputs or outputs, followed by two open-ended questions (Supplementary Material 1). The initial 41 items were rated on a 4-point scale (−2 strongly disagree; −1 moderately disagree; +1 moderately agree; +2 strongly agree); all items also had the options IDK (I don't know) and NA (not applicable). The 41 items were assigned to the following topical categories: Activity Patterns, Climate Range, Enclosure Complexity, Diet, Endocrinology, Enrichment, Keeper Interactions, Keeper Observations, Physical Appearance, Animal Safety, Sensory Environment, Social Setting, Surfaces, Training Program, Visitor Interactions, and Weight. The categories were intended more to support a discussion process than to strictly characterize items. The final two open-ended questions asked raters to suggest three changes they would recommend to improve the welfare of the animals and three aspects of the animals' current care they believe promoted good welfare. The open-ended items were included to facilitate quality discussion and action planning to enhance animal care and welfare.

2.2. Welfare discussion tool protocol

The WDT was completed for each species in our care minimally once per calendar year. The WDT was completed by a team of three people: (1) a Curator or Zoological Manager responsible for that animal, (2) an Animal Caretaker from the animal's area, and (3) a person that is not responsible for the daily care of that individual animal but has taxonomic expertise with the species being evaluated. The outside person may be any zoo employee who is familiar with the species' natural and life history (e.g., scientist, animal care professional from another area of the zoo).

For animals that were individually identifiable (164 species), raters were asked to consider each individual animal independently when completing the WDT for that species. For each statement, the rater provided one rating per individual animal. For socially-housed animals that were not individually identifiable and did not have individual accession numbers (10 species, e.g., Madagascar hissing cockroaches; desert walking sticks; African cichlids) a Group WDT was used. For the Group WDT, input item ratings were applied to the full group since care generally didn't differ between individuals within a group, and for output items, raters reported the percentage of the group to which they assigned each rating. For example, for a given item a rater may provide a + 2 rating to 80% of the group and a + 1 rating to 20% of the group. A copy of the Group WDT is provided in Supplementary Material 2.

Each rater completed the WDT independently within two weeks prior to meeting in person to discuss their ratings and generate a list of action items intended to enhance animal welfare. All ratings and associated action items were entered into Lincoln Park Zoo's animal records software Tracks® (Tracks Data Solutions, 2023) using Tracks' Welfare Module.

2.3. Data collection and analysis

Data for this study were extracted from Tracks®. We considered all WDT ratings completed between the start of the zoo-wide animal welfare assessment program in July 2019 and December 2021 and compared scores across the following taxonomic groups: mammals, birds, reptiles, amphibians, fish, and invertebrates. We were also interested in whether scores differed across taxonomic groups for items classified as inputs and/or outputs, as these might provide insight as to whether there were differences in care provided or our interpretation of the animals' responses to that care, respectively.

We also wanted to better understand raters' confidence in judging the welfare of different taxonomic groups and raters' assessments that the statements were relevant for the species they were considering. For that reason, we compared the count of NA (not applicable) and IDK (I don't know) ratings across taxonomic groups. NA scores and IDK scores were analyzed collectively because these scores were not differentiated in Tracks® (herein referred to as NA/IDK scores) and because we believe that raters may have used NA and IDK interchangeably when assessing welfare.

Average WDT scores for each species, and corresponding taxonomic groups, were calculated in Microsoft Excel. Statistical analyses were performed in R version 4.2.1 (R Core Team, 2017). To account for variation in the total count of WDT ratings between species (differences in number of raters and assessment events), species' scores were averaged so that each species accounted for one data point in the dataset. Two-way ANOVAs were used to test if taxonomic group and type of WDT (Individual or Group) predicted species' average WDT scores, species' average input scores, species' average output scores, and the total count of NA/IDK scores per species. Separate two-way ANOVAs were run for each response variable. Assumptions for ANOVAs were met prior to analysis (data were normally distributed, independent, and homogeneity of variance was not violated). Post-hoc Tukey HSD tests were used to determine which taxonomic groups significantly differed from each other. Finally, we calculated average scores for each taxonomic group for each topical category (Enclosure Complexity, Physical Appearance, etc.) to better characterize our data, but did not run formal analysis on these groupings of items.

3. Results

There were 939 Welfare Discussion Tools completed by 115 people for 174 species between July 2019 and – December 2021. This included 432 mammals, 261 birds, 182 reptiles, 37 amphibians, 14 fish and 13 invertebrates. Table 1 indicates the mean WDT score for each taxonomic group for each of the individual topical categories.

A two-way ANOVA revealed that WDT scores differed significantly by tool type ($F(1) = 4.01$, $P = 0.05$) and taxonomic group ($F(5) = 5.53$, $P < 0.01$). Animals rated using the Individual WDT scored higher than animals rated using the Group WDT ($P = 0.05$). Post-hoc Tukey HSD tests revealed that the WDT scores for the taxonomic group mammal were significantly higher than scores for bird and fish ($P < 0.01$ and 0.05, respectively) (Fig. 1).

When considering mean WDT scores in terms of inputs only (resource-based measures), a two-way ANOVA revealed that average input scores differed significantly by tool type ($F(1) = 4.08$, $P = 0.05$) and taxonomic group ($F(5) = 4.61$, $P < 0.01$). Animals rated using the Individual WDT scored higher than animals rated using the Group WDT ($P = 0.05$). Post-hoc Tukey HSD tests revealed that taxonomic group mammal had significantly higher input scores than bird ($P < 0.01$ (Fig. 2).

When considering mean WDT scores in terms of outputs only (animal-based measures), a two-way ANOVA revealed that average output scores differed significantly by taxonomic group ($F(5) = 4.65$, $P < 0.01$). Average output scores did not significantly differ by tool type ($F(1) = 2.21$, $P = 0.14$). Post-hoc Tukey HSD tests revealed that taxonomic group mammal had significantly higher output scores than bird and fish ($P = 0.01$ and $P = 0.05$, respectively). (Fig. 3).

A two-way ANOVA revealed that the total count of NA/IDK scores differed significantly by taxonomic group ($F(5) = 5.99$, $P < 0.01$). On average, NA/IDK was chosen the least for mammals (4.82) and was chosen the most for invertebrates (22.33); the mean number of times NA/IDK was chosen for other taxonomic groups is as follows (Bird: 6.98, Reptile: 10.98, Fish: 11.46, Amphibian: 11.78). Post-hoc Tukey HSD tests revealed that taxonomic group mammal had significantly fewer NA/IDK scores than bird and reptile ($P < 0.01$ and $P = 0.03$, respectively).

4. Discussion

Our primary aim was to determine whether a taxonomic difference emerged in the scores of our species-general welfare tool. Taxonomic differences were clearly present in the Welfare Discussion Tool scores analyzed from the first two and a half years of data collected at Lincoln Park Zoo. The rating scales ranged from –2 (presumably poor welfare) to + 2 (presumably positive welfare) and all taxonomic groups scored positively, on average. Specifically, mammals scored higher than birds and fish in their overall WDT score. Additionally, mammals scored higher than birds in their input scores and higher than birds and fish in their output scores. Mammals also received fewer ratings of "Not applicable" and/or "I don't know," indicating greater knowledge or confidence by raters in applying the tool to assess the welfare of mammals.

There are multiple possible explanations for why mammals scored higher on the WDT relative to other taxonomic groups. Perhaps the simplest explanation is that the tool itself was biased toward mammals in a way that supported more positive scores. While the intention was to create a species-general welfare assessment tool and people and resources (Kagan et al., 2015) with broad taxonomic expertise were consulted, it is possible that the tool authors' general familiarity and

Table 1
Mean WDT for each taxonomic group for each of the individual topical categories. The highest score in each category is italicized and bolded. *Note regarding endocrinology: data on endocrinology are not available for many species and data should be interpreted cautiously.

Topical Category	Mammal	Bird	Reptile	Amphibian	Fish	Invertebrate
Activity Patterns	1.30	***1.43***	1.02	0.76	1.06	0.94
Climate Range	0.56	-0.13	***0.81***	0.33	0.40	-0.61
Enclosure Complexity	0.82	0.95	1.00	0.98	0.56	***1.01***
Diet	0.80	0.42	0.38	1.03	-0.01	***1.17***
Endocrinology	1.07	0.86	***2.00***	1.00	NA	NA
Enrichment	***1.09***	0.36	-0.27	0.12	-0.56	0.25
Keeper Interactions	***1.22***	0.79	0.74	0.60	0.62	1.05
Keeper Observations	***1.45***	1.30	1.36	1.43	0.98	1.12
Physical Appearance	1.04	0.85	0.95	1.09	0.63	***1.58***
Animal Safety	1.03	0.56	***1.55***	1.42	1.52	0.54
Sensory Environment	0.71	0.58	0.77	0.55	***0.86***	0.50
Social Setting	0.88	0.62	***0.98***	0.77	0.53	0.91
Surfaces	0.68	***1.02***	0.50	0.32	0.13	0.99
Training Program	***1.03***	-0.20	-0.98	-1.27	-1.61	NA
Visitor Interactions	***0.78***	0.43	0.75	0.45	0.55	0.56
Weight	1.14	0.27	***1.24***	0.91	-1.88	NA

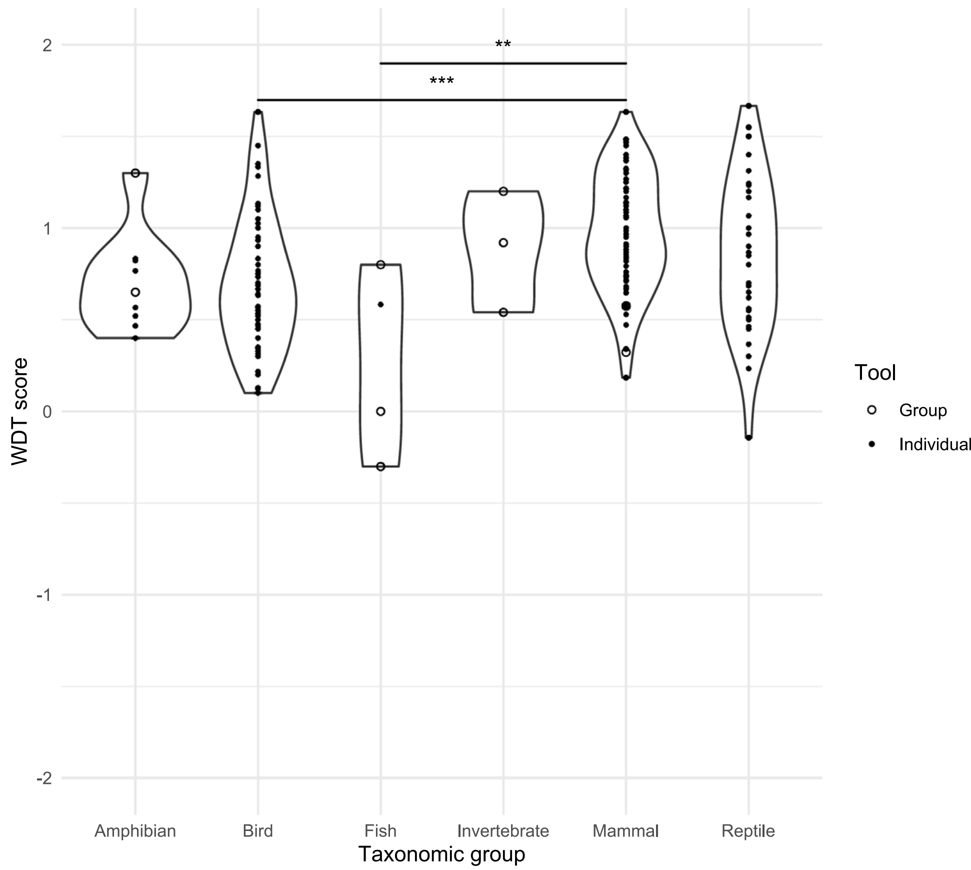

Fig. 1. Violin plot of average WDT scores for each taxonomic group. Each dot represents one species and species assessed with the Group WDT are shown as open dots. Taxonomic groups with significantly different WDT scores are connected by a line. Significant differences corresponding to $P < 0.05$ are indicated with a single asterisk and differences of $P < 0.01$ are indicated with double asterisks.

experience with mammals, informed in part by the greater availability of scientific knowledge related to mammal welfare (Binding et al., 2020) introduced unintentional biases in the phrasing of statements in the WDT in a way that set the stage for better scores for mammalian species.

An additional possibility is that high levels of caretaker interaction with the mammal species at Lincoln Park Zoo may have unintentionally biased raters to rate the welfare of these animals higher than other taxonomic groups. The topical categories that mammals scored highest in (Enrichment, Keeper Interactions, Keeper Observations, Training Program, and Visitor Interactions) all involve caretaker interaction, which may introduce biases in welfare assessments. High levels of caretaker interaction may cause some raters to develop greater connections with individual animals, which may lead to more optimistic judgments during welfare assessments (Thomas et al., 2022; Hosey and Melfi, 2012).

Yet another possibility has to do with the animals themselves. Mammals have a behavioral repertoire that is both generally familiar and easily observable by humans. It is possible that our general familiarity with the activity levels and behavioral repertoires of mammals, and conversely, our unfamiliarity with the activity levels and behavioral repertoires of other taxa, unintentionally favorably bias our understanding and assessments of their welfare (see also De Waal 1999; Harrison and Hall, 2010; Serpell 2019 for related discussions of how phylogenetic similarity influence our perceptions of other animals). Notably, NA/IDK was chosen, on average, least on WDT reports for mammals relative to other taxonomic groups, indicating that raters were able to provide a numerical score for more of the statements on the WDT when assessing mammals than when assessing other taxa. While this tells us that there are meaningful differences in raters' knowledge of the mammals' care, their knowledge of the mammals' responses to that care, or their judgment of the relevance of the item for mammals' welfare (all relative to other taxonomic groups), this does not help us determine whether the mammal bias emerges from a bias in the WDT itself or a bias in the raters.

A final possibility, not mutually exclusive to those above, is that mammals are actually experiencing better welfare than some other taxonomic groups. Considered in the broader context of the known taxonomic biases in the fields of animal behavior and animal welfare research (Binding et al., 2020; Rosenthal et al., 2017), it is possible that greater general knowledge of mammalian species has facilitated more species-appropriate human care for these animals that has resulted in enhanced welfare. Of note, mammals scored higher in both inputs and outputs, indicating that raters judge the care we are providing and the responses the animals are showing us to both indicate positive welfare. However, at this time, it is not possible to determine whether mammals are actually experiencing better welfare relative to other taxonomic groups or if aforementioned biases lead to more positive welfare ratings for mammalian species, independent of the actual welfare status of the animals.

Tool type was a significant predictor of overall welfare scores and input (but not output) scores. Specifically, these scores were lower on the Group WDT compared to the Individual WDT. We interpret this cautiously considering the small number of species assessed with the Group WDT and that the Group WDT was rarely applied to mammalian species (Figs. 1–3). We think this effect may have been driven by differences in the raters' comfort assessing welfare of groups, particularly since there are only a small number of species assessed with the Group WDT. Notably, NA/IDK was chosen the most for invertebrates compared to other taxonomic groups, indicating that raters were not able to provide a numerical score for many of the statements on the Group WDT

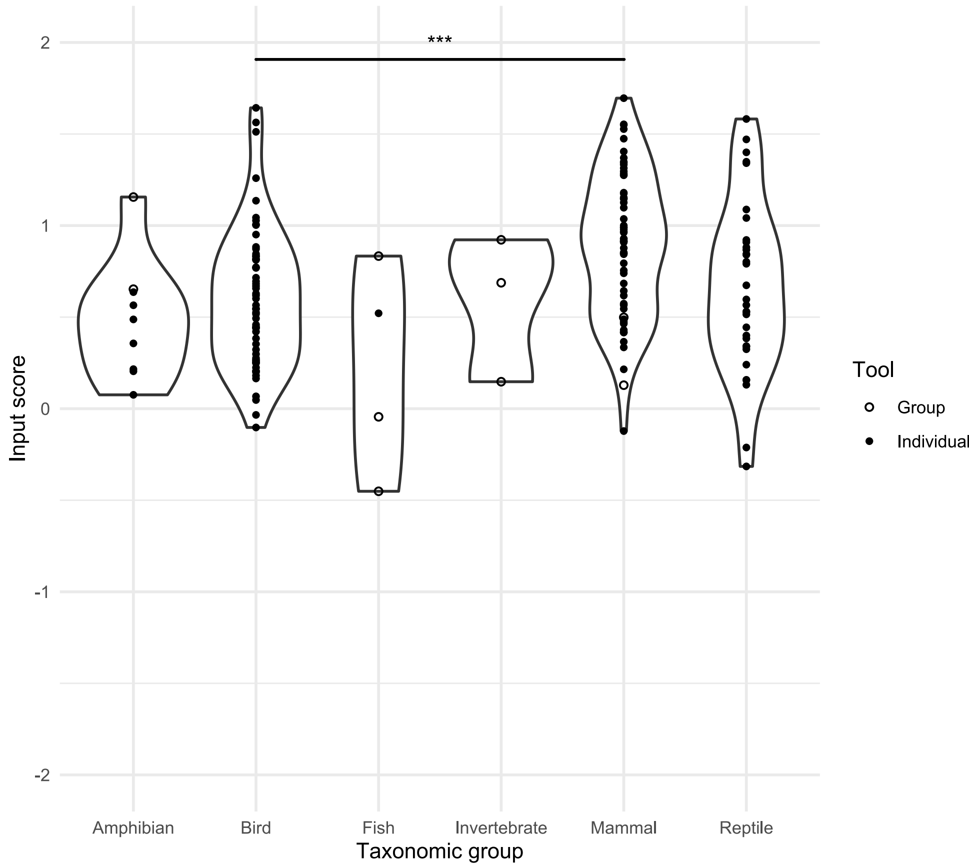

Fig. 2. Violin plot of average input scores for each taxonomic group. Each dot represents one species and species assessed with the Group WDT are shown as open dots. Taxonomic groups with significantly different input scores are noted. Significant differences corresponding to P < 0.05 are indicated with a single asterisk and differences of P < 0.01 are indicated with double asterisks.

when assessing invertebrates.

Based in large part on the findings presented in this study, we revised the Individual WDT and Group WDT (Supplementary Material 3 and 4). With input from animal welfare scientists, animal caretakers, behavioral husbandry experts and veterinary staff, we focused on three approaches to broaden the taxonomic relevance of the tool and minimize potential rater biases. First, we identified opportunities for item revisions to be more taxonomically inclusive. For example, the following items were revised to be more relevant to non-mammalian species: Original item: *(Input) There is a training program to facilitate voluntary shifting between areas (*e.g.*, movement between exhibit and holding).*

Revised item: *(Input) There is training to facilitate voluntary shifting between areas (*e.g.*, between exhibit and holding, into transfer tube or secondary area during cleaning, etc.). Only in the case where the animal rarely to never moves between spaces in their lifetime should NA be chosen.*

Original item: *(Output) The animal's behavior indicates comfort with sensory input (visual, auditory) from visitors.*

Revised item: *(Output) The animal's behavior, use of space, and/or posture indicates comfort with sensory input (visual, auditory, olfactory, gustatory, tactile) not negatively impacted by visitors and staff.*

Second, we added new items with broad species relevance that emerged in the action plans for many non-mammalian species, for example, *The exhibit has UV lighting consistent with the species natural habitat.*

Third, we wanted to eliminate the possibility that raters chose "not applicable" on items that could be informative. We thought that removing many NA opportunities may push raters to consider whether an item was relevant to the species in question and encourage less mammal-centric bias in their evaluation. In the original version of the WDT, NA was available on all 41 items, on the revised version of the tool NA is available on 6 items, and in these instances, clarifications have been provided for when "NA" should be chosen. Raters can still choose "I don't know" on all items. We are planning for a similar retrospective analysis of the updated tool in the coming years to determine how these revisions have impacted the taxonomic patterns reported here.

Collectively, the results presented in this study yet again highlight the importance of focusing future animal behavior and animal welfare research on a broader taxonomic range, with the specific goal of learning more about species that have been historically understudied. Zoos and aquariums care for a wide variety of taxa yet research in zoos and aquariums is still largely focused on endothermic vertebrates. These taxonomic biases in research foci and consequent species knowledge may, unintentionally, contribute to biases in animal welfare measurement. Knowledge gained from research focused on more-often studied species may be applied to other taxa, unintentionally resulting in suboptimal care for those animals. For some species, so little may be known that starting with basic research into activity budgets, space use, and behavioral repertoires may generate important knowledge that can be built upon to more effectively care for and assess welfare of those species.

5. Conclusion

Evidence-based care should be the industry standard in zoos and aquariums (Melfi, 2009). As institutions continue to advance on this path, it is important to consider which outputs can reliably inform us about the welfare of non-mammalian taxa. Well-designed studies of space use and activity budgets for traditionally understudied taxa can greatly improve the knowledge base, and consequently the care and welfare, of these species. Importantly, as knowledge about understudied

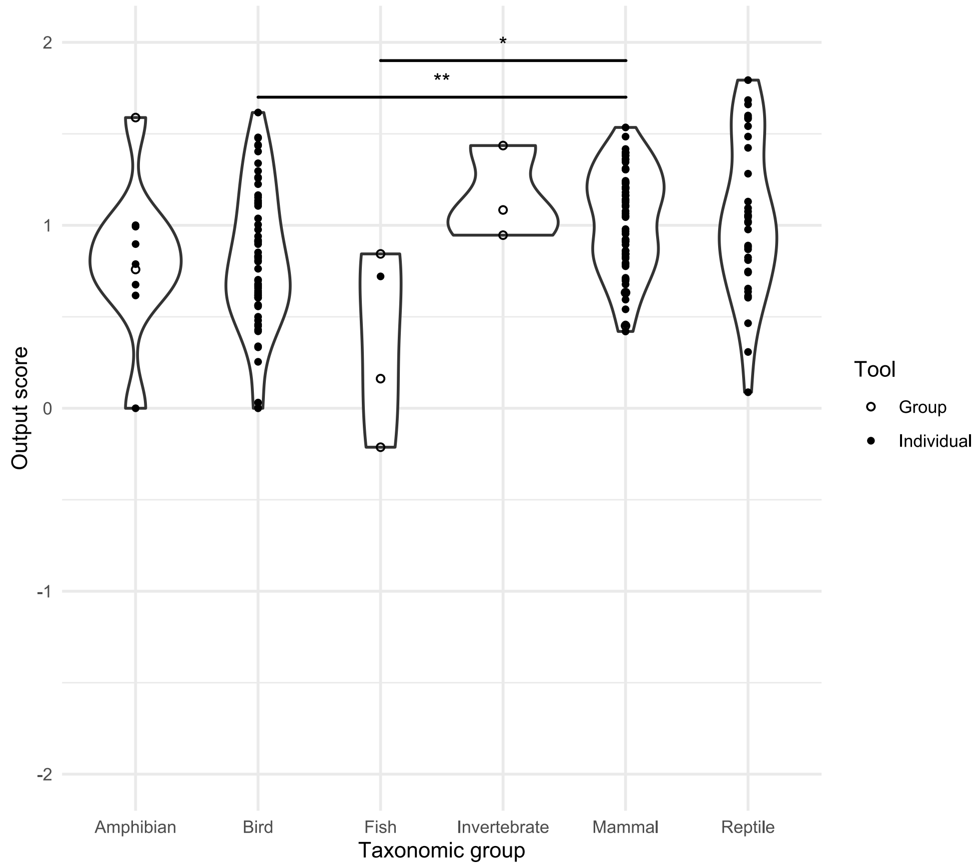

Fig. 3. Violin plot of average output scores for each taxonomic group. Each dot represents one species and species assessed with the Group WDT are shown as open dots. Taxonomic groups with significantly different output scores are noted. Significant differences corresponding to $P < 0.05$ are indicated with a single asterisk and differences of $P < 0.01$ are indicated with double asterisks.

species grows, welfare assessments should be updated to ensure that animal welfare is being accurately assessed and that the animals are provided with the best possible care. We invite other institutions interested in species-general welfare assessments to consider adopting a similar internal assessment approach and/or potentially adopting a version of our improved welfare discussion tool to serve their welfare assessment needs.

Declaration of Competing Interest

This work has not been published before and is not under consideration at another journal, book, or other outlets. All authors have approved this version of the manuscript and declare no conflicts of interest.

Acknowledgements

We would like to thank Owen Howard, DePaul University's Dean's Undergraduate Fellow, whose internship project inspired this study and Vanessa Godbolt for entering WDT data into Tracks®. We'd also like to thank the Animal Care staff at Lincoln Park Zoo, including Maureen Leahy, Dave Bernier, Steve Ross, Adrienne Horrigan, Allison Kao, Amanda Barnes, Jason Wark, and Natasha Wierzal for their role in developing and updating the tool and for their continued inspiration and support. Funding for the publication of this study was provided by the American Society for the Prevention of Cruelty to Animals® (ASPCA®) Open-Access Publishing Fund.

Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at doi:10.1016/j.applanim.2023.105965.

References

Association of Zoos and Aquariums, 2023, The Accreditation Standards and Related Policies. https://assets.speakcdn.com/assets/2332/aza-accreditation-standards.pdf, accessed online 24 April 2023.
Benn, A.L., McLelland, D.J., Whittaker, A.L., 2019. A review of welfare assessment methods in reptiles, and preliminary application of the welfare quality® protocol to the pygmy blue-tongue skink, *Tiliqua adelaidensis*, using animal-based measures. Animals 9 (1), 27.
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.
Brambell, R., 1965. Report of the Technical Committee to Enquire into the Welfare of Animals Kept under Intensive Livestock Husbandry Systems. Cmd. (Great Britain. Parliament). H.M. Stationery Office,, London, UK, pp. 1–84.
Clegg, I.L., Borger-Turner, J.L., Eskelinen, H.C., 2015. C-Well: The development of a welfare assessment index for captive bottlenose dolphins (*Tursiops truncatus*). Anim. Welf. 24 (3), 267–282.
R. Core Team, 2017, R: A language and environment for statistical computing. R foundation for statistical computing, Vienna, Austria. URL. https://www.R-project.org/.
Dawkins, M.S., 2021. The Science of Animal Welfare: Understanding What Animals Want. Oxford University Press,, USA.
De Waal, F., 1999. Anthropomorphism and anthropodenial: consistency in our thinking about humans and other animals. Philos. Top. 27 (1), 255–280.
Greggor, A.L., Vicino, G.A., Swaisgood, R.R., Fidgett, A., Brenner, D., Kinney, M.E., Lamberski, N., 2018. Animal welfare in conservation breeding: applications and challenges. Front. Vet. Sci. 5, 323.
Harrison, M.A., Hall, A.E., 2010. Anthropomorphism, empathy, and perceived communicative ability vary with phylogenetic relatedness to humans. J. Soc., Evolut., Cult. Psychol. 4 (1), 34.

Hill, S.P., Broom, D.M., 2009. Measuring zoo animal welfare: theory and practice. Zoo. Biol.: Publ. Affil. Am. Zoo. Aquar. Assoc. 28 (6), 531–544.

Holst, M.M., Miller-Morgan, T., 2021. The Use of a species-specific health and welfare assessment tool for the giant pacific octopus, *Enteroctopus dofleini*. J. Appl. Anim. Welf. Sci. 24 (3), 272–291.

Honess, P., Wolfensohn, S., 2010. The extended welfare assessment grid: A matrix for the assessment of welfare and cumulative suffering in experimental animals. Altern. Lab. Anim. 38 (3), 205–212.

Hosey, G., Melfi, V., 2012. Human–animal bonds between zoo professionals and the animals in their care. Zoo. Biol. 31 (1), 13–26.

Jones, N., Sherwen, S.L., Robbins, R., McLelland, D.J., Whittaker, A.L., 2022. Welfare assessment tools in zoos: from theory to practice. Vet. Sci. 9 (4), 170.

Kagan, R., Carter, S., Allard, S., 2015. A universal animal welfare framework for zoos. *J. Appl. Anim. Welf. Sci.* 18 (sup1), S1–S10.

Melfi, V.A., 2009. There are big gaps in our knowledge, and thus approach, to zoo animal welfare: a case for evidence-based zoo animal management. Zoo. Biol.: Publ. Affil. Am. Zoo. Aquar. Assoc. 28 (6), 574–588.

Mellor, D.J., 2017. Operational details of the five domains model and its key applications to the assessment and management of animal welfare. Animals 7 (8), 60.

Rose, P., O'Brien, M., 2020. Welfare assessment for captive Anseriformes: A guide for practitioners and animal keepers. Animals 10 (7), 1132.

Rosenthal, M.F., Gertler, M., Hamilton, A.D., Prasad, S., Andrade, M.C., 2017. Taxonomic bias in animal behaviour publications. Anim. Behav. 127, 83–89.

Serpell, J.A., 2019. How happy is your pet? The problem of subjectivity in the assessment of companion animal welfare. Anim. Welf. 28 (1), 57–66.

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.

Temple, D., Courboulay, V., Manteca, X., Velarde, A., Dalmau, A., 2012. The welfare of growing pigs in five different production systems: assessment of feeding and housing. Animal 6 (4), 656–667.

Thomas, V., Orwig, J., Shelmidine, N., York, J., 2022. Zookeeper–Animal Bonds and Their Relationship with Conservation Action. J. Zool. Bot. Gard. 3 (4), 699–713.

Tracks Data Solutions | Tracks Software, 2023, Online: trackssoftware.com.

Veasey, J.S., 2017. In pursuit of peak animal welfare; the need to prioritize the meaningful over the measurable. Zoo. Biol. 36 (6), 413–425.

Whitham, J.C., Wielebnowski, N., 2013. New directions for zoo animal welfare science. Appl. Anim. Behav. Sci. 147 (3–4), 247–260.

Wolfensohn, S., Shotton, J., Bowley, H., Davies, S., Thompson, S., Justice, W.S., 2018. Assessment of welfare in zoo animals: Towards optimum quality of life. Animals 8 (7), 110.

Zoo and Aquarium Association Australasia, 2023, 8.1 Policy – Animal Welfare. https://zooaquarium.org.au/common/Uploaded%20files/Policies/P_Animal%20Welfare.pdf. Accessed online 24 April 2023.