

Article

# Investigating Personality in Polar Bears Using the Concept of Behavioral Instability

Anika Gottschalk<sup>1†</sup>, Anne Cathrine Linder<sup>1†</sup>, Henriette Lyhne<sup>1†</sup>, Marie Gade Langbak<sup>1†</sup>, Trine Hammer Jensen<sup>2</sup> and Cino Pertoldi<sup>1,2\*</sup>

<sup>1</sup> Department of Chemistry and Bioscience, Aalborg University, Fredrik Bajers Vej 7H, Aalborg DK-9000, Denmark

<sup>2</sup> Department of Zoology, Aalborg Zoo, Mølleparkvej 63, Aalborg DK-9000, Denmark

†All four authors contributed equally

\* Correspondence: cp@bio.aau.dk

**Abstract:** Behavioral instability is a new concept used for indicating environmental stress based on behavioral traits. This study investigates the possibility of using behavioral instability as a tool for assessing personality in captive animals. The understanding of personality in captive animals can be a resourceful tool in the development of enrichment programs in order to improve animal welfare. In this study it is examined how an olfactory stimulus affects the behavior of two individuals of the species *Ursus maritimus* in captivity. When using continuous focal sampling throughout the day it was for many behaviors found that the individuals responded differently to stimuli, indicating that there was a difference in personality. This is shown using multiple approaches. One approach used traditional methods for behavioral analyses and the other approach used the concept of behavioral instability as a new quantitative method. This study proves use of behavioral instability as a new quantitative method for investigating personality, expanding the possibility of comparing personality between species. Moreover, it is shown that outliers, which cause asymmetric distributions, should not be removed in behavioral analysis without careful consideration.

Keywords: asymmetry; ethogram; olfactory stimuli; stereotypy; *Ursus maritimus*; enrichment; captivity; asymmetric diversity

## 1. Introduction

It has for several species been shown that conspecific individuals have different personalities [1, 2]. These different personalities are expressed by consistent behavioral responses under various conditions, that can vary with different variables e.g. population density, stress and enrichment [2, 3, 4]. The personality of an animal can influence its welfare, as these responses can vary between individuals, i.e. an environmental condition may be well tolerated by one individual, but not by another [5]. Stereotypic behavior is described as a repetitive motion with no apparent purpose and has generally been shown to be a sign of stress, due to its correlation with increased corticoid levels, thus making stereotypy an indication of poor welfare [6, 7]. Mason and Latham [7] investigated situations that caused an increase in stereotypic behavior and found that welfare decreased significantly in 68% of these situations. Pacing (stereotypic locomotion) is an example of stereotypic behavior, it often begins as anticipatory pacing, developing into generalized stereotypic behavior when the animal experiences stress [6, 8]. Stereotypic behavior is especially observed in predatory animals, moreover it has been found that bears display a higher frequency of stereotypic behavior than canids and felids [6]. This could partly be explained by their large home ranges in the wilderness, along with the accessible space in captivity only being a small proportion of the space they inhabit in the wilderness [6, 8, 9]. This explains the tendency of bears to pace near the limits of their enclosure, where pacing can represent an urge to explore more than the limits of the enclosure allow [6, 10]. Ross [11] studied how access to an indoor enclosure affects the stereotypic behavior of polar bears and showed that they become less stereotypic throughout the day, even though they only spent a small proportion of time indoors. This indicates that the ability to choose made them less stereotypic rather than the physical act of being inside [11]. This could also apply to the presence of enrichment, as shown for olfactory stimuli by Carlstead and Seidensticker [12], where a black bear throughout the day spent less time on stereotypic behavior and more time exploring the yard, when it had been given enrichment in the form of odor from

other bears. They concluded that an olfactory stimulus, at least during breeding season, was sufficient to distract the bear from pacing [12].

Several studies have shown that enrichment and the presence of choice in activity is negatively correlated with stereotypy [9, 11, 13]. However, other studies have shown that not all enrichments improve welfare when measured in time spent on stereotypy [12], this could be explained by the variation in the tested individuals' personalities [5]. Personality has been found in different bears [1], but only few studies exist for polar bears, most of them studying stereotypic behavior in general [10, 11, 14]. To improve the welfare of polar bears and other large predatory animals in captivity, it would be relevant to quantify their behavior and personality in order to understand how their general welfare and the welfare of each individual can be improved [8]. It is therefore important to investigate if different animals have different personalities, as they would be expected to react differently to stimuli, either increasing or decreasing their time spent on stereotypic behavior leading to a difference in welfare as well. Rose et al. [8] and Shyne [13] emphasize the need for further development of quantitative assessments of animal welfare in order to increase the reliability of non-invasive welfare indicators, such as behavioral traits.

The sampling methods used in the traditional assays of animal behavior vary between studies and have been described and compared in Altmann [15]. Many of these studies used an interval instantaneous sampling technique, where the behavior is observed at a specific time and then noted as state behavior for the entire time interval between the current and previous observation, assuming that the individuals were performing this behavior the whole time [15, 16, 17]. Other studies only sample from the beginning of the behavioral state they wish to analyze until the animal switches to another behavioral state, thus only observing a single behavior at a time [10, 14]. In another study, focal sampling of a shorter period was used [17]. Each of these sampling techniques can be used for different analyses, answering different questions [15]. The time intervals, at which observations sessions took place, also varied greatly between studies [11, 16, 17]. Bashaw et al. [17] found that there was a difference in behavior throughout the day, proving that the assessment should be carried out not only at a specific time of the day, but for a longer period of time, covering a larger percentage of the day. Standardizing these sampling methods would contribute to a quantitative and systematic behavior analysis.

There have been made different suggestions for improving the traditional non-standardized method using ethograms and observations of different time intervals, by using more quantitative and systematic methods. Pertoldi et al. [18] introduced the concept of behavioral instability based on the concept of developmental instability. Behavioral instability was introduced as a method of studying the symmetry of behavior, by observing bilateral behavioral traits e.g. how many times an individual looks to the left versus the right. Bech-Hansen et al. [19] introduced two variables to this concept, BSYM and BVAR. BSYM is the behavioral instability of symmetry, meaning the deviation from a symmetric distribution for the studied behavior, and BVAR is the variance of residuals for the studied behavior, where a higher variance indicates a smaller capacity for anticipating a behavior when stressors are present [19]. The concept of behavioral instability could, as proposed by Bech-Hansen et al. [19], also be applied to measure the effect of environmental stress on behavioral data other than bilateral data as it can be used to measure the effect of environmental stress. Therefore, it should be investigated, how behavioral instability can be used as a new, quantitative way of studying behavior and, moreover, personality.

## **Aim of the paper**

This study investigates the application of the concept of behavioral instability as a tool for studying the personality of captive animals. This was done by investigating the personality of polar bears in captivity by comparing the effect of olfactory stimuli on two individuals at Aalborg Zoo, Denmark. It was expected that the stimuli would have an effect on the individuals and their behavior and furthermore that there would be a difference in the two individuals' personalities, which would be indicated by a difference in how the individuals reacted to the stimuli.

## **2. METHODS**

### *2.1. Participants and Setting*

In this study the behavior of two female polar bears at Aalborg Zoo in Denmark was observed. The two individuals are siblings that were born in November 2016 at Aalborg Zoo. The sisters have since spring 2019 been kept in a separate enclosure from their mother. The two enclosures were separated by a dry moat, giving the two individuals visual access to their mother. Their diet consisted of vegetables, fruit, fish, meat (primarily horse intestines), dog kibble and various treats such as dried dates, which they were fed randomly throughout the week. The area of the enclosure used for this study was 768 m<sup>2</sup> and consisted of a pool, land covered by gravel and concrete, and a den (map of the enclosure can be seen in Appendix H). The windows for the zoo visitors were placed opposite the den, making the inside of the den not visible for visitors. The zookeepers were able to access the polar bears when they were in the den, this is also where the zookeepers occasionally would train the polar bears and feed them treats.

## 2.2. Data collection

The observations took place from primo October to primo November 2019 during the zoos off-season. Nine observation sessions were spread throughout this time period. The observations sessions were conducted by filming the polar bears using four action cameras (Kitvision Escape HD 5), that were placed around the enclosure, ensuring video surveillance of the entire outdoor perimeter (camera placement can be seen in Appendix H). Each session began at sunrise, ranging from 07:29 am (UTC+2) to 08:34 am (UTC+1) and lasted for nine hours. Three of the observation sessions were control treatments (treatment C), that were used as a baseline measurement of the polar bears behavior under normal conditions. During three of the other observation sessions the bears were given stimuli in the form of two dog scented objects (treatment D), one for each individual, which were thrown into the enclosure between 09:00 and 09:30 am (the object can be seen in Appendix I). The objects were fabric boxes that each were in contact with a different dog for approximately a week prior to each of the three observation sessions, thus scenting the boxes with the natural odor of the dogs. In order to estimate the effect of the dog odor and not the effect of the object itself, three observation sessions were used to observe the effect of the unscented fabric boxes. The behavioral data for the observation sessions with unscented objects was only used to confirm that the effect of the stimuli came from the dog odors and not the fabric boxes themselves and this data was therefore only used in a preliminary analysis (Appendix A and F.1-3).

## 2.3. Analysis

Behavioral observations were based on the analysis of the filmed material by four coders, using the ethogram described in table 1. Prior to this a concordance test was performed to ensure that the inspections of all four coders were in agreement. The footage was analyzed using continuous focal sampling of the nine hours that each observation session lasted [15]. Furthermore, all occurrences were treated as states as described by Altmann [15], thus for each observation session all 32,400 seconds were coded. The preliminary analysis was based on all nine observation sessions, amounting to 291,600 seconds and 3,322 data points. Further analyses were based on only six observation sessions, three for treatment C and three for treatment D, amounting to 194,400 seconds and 2,236 data points.

The statistical analyses were conducted in RStudio version 3.6.0 [20] and Past version 3.26b [21]. As the data was not normally distributed, outliers were removed by two different methods. This resulted in three versions of the data set: one containing all the original data points, one with only data points inside the interquartile range (IQR), thus removing all data points outside the interval between the 25th and 75th percentile and one with outliers removed using Median Absolute Deviation method (MAD) with the conservative threshold value of 3 [22]. The data was analyzed using non-parametric methods for all three data sets. Most analyses were conducted for each observation session separately, due to some significant differences between the observation sessions of the same treatment. However, some analyses were only possible to conduct for the data in which all three observation sessions were pooled for each treatment and each individual separately.

The percentage of time each individual used on each behavior was estimated for the different observation sessions in order to examine the differences in the distribution of time spent on each behavior both between treatments and individuals. Furthermore,  $\chi^2$  tests with Yates correction [23] were carried out on pooled data, with the variables being the different treatments and the two individuals. This was only done for the data set containing all data points, as it was only for this data set that all nine hours were represented.

For all data sets the medians, variances, asymmetry indices (skewness) and kurtoses were calculated to examine the differences in time each behavior lasted per occurrence, how much it varied and the shape of the data between individuals and treatments. Due to the non-normal distribution of the data, the variances were based on the IQR. The differences in medians and variances were tested with pairwise comparisons using the Mann-Whitney U-test and Levene's test [23] for all data sets, respectively. Pairwise comparisons of the asymmetry indices and kurtoses were made for all data sets by comparing the bootstrap confidence intervals of the observation sessions. For the pairwise comparisons a significance level of 0.05 was used. Additionally, correlations of the time each behavior lasted per occurrence were calculated between the different treatments as well as between the two individuals using Pearson product-moment correlation [23] (Appendix K and L).

In order to observe the distribution of time spent on each behavior throughout the day, cumulative curves were plotted for each individual and observation session (Appendix C). For the cumulative data Pearson product-moment correlations between the different individuals and treatments were calculated based on the time of the day at which the individual completed a certain percentage of the total amount of time spent on a given behavior (i.e. T10, T20, ..., T100) (Appendix J). The cumulative curves and correlations between them were only calculated for the data set containing all data points.

Medians were calculated for each treatment and individual using the pooled data from the data set containing all data points and the MAD data set. For each behavior, the medians for both individuals and treatments were plotted along with a trend line between the median of treatment C and median of treatment D for each individual. The slopes of the trend lines were calculated as well as the percentage differences in the trend line slopes between the two individuals for the same behavior.  $\chi^2$  tests with Yates correction were carried out comparing these medians between the two treatments and the two individuals for each behavior (Appendix D). This procedure was also carried out for the variances, asymmetry indices and kurtoses of the pooled data for the data set containing all data points. The same plots were made for the two data sets where outliers had been removed (Appendix E).  $\chi^2$  tests were also executed comparing the variances, asymmetry indices and kurtoses, respectively (Appendix D).

### 3. Results

The time spent on different behaviors varied between all the observation sessions and the individuals. Figure 1 shows that individual 2 generally spent a larger amount of time on stereotypic behavior and a smaller amount of time on inactivity compared to that spent by individual 1. However, the amount of time the two individuals spent on these behaviors varied greatly between the observation sessions. When comparing the two individuals stereotypic and inactive behavior for treatment D respectively, a significant difference between the two individuals was observed for both behaviors. For this treatment individual 1 spent a greater amount of time being inactive than individual 2, whereas the opposite was found for the amount of time the individuals spent on stereotypic behavior. Furthermore, it was found that individual 1 spent a significantly greater amount of time being inactive for treatment D in comparison to treatment C.

When comparing the three observation sessions of treatment C with those of treatment D it was generally found for both individuals that the median and variance of the amount of time spent on activity both on land and in water were greater for treatment D than treatment C (table 2). The same tendency applies for both individuals when comparing inactive behavior between treatment C and treatment D. In contrast to this it was found that the median time and variance in time spent on social play were generally greater for treatment C than treatment D for both individuals. In regard to stereotypic behavior it was for individual 1 found that the median time and variation in time were greater for treatment D than treatment C, there was however also found a few significant results showing these variables as being greater for treatment C than for treatment D. Contrary to this, it was for individual 2 found that the median times spent on stereotypic behavior only showed significant differences where the median of treatment D was greater than that of treatment C. This was also roughly the case when comparing the variance in time spent by individual 2 on stereotypic behavior. In this case all but a single comparison in the MAD data set showed the variance of treatment D to be significantly greater than that of treatment C.

The significant results observed for the comparison of the medians and variances of the two treatments varied slightly between the three data sets, but the results of all three data sets generally showed the same tendencies (table

2). The slight variation occurred because the two data sets IQR and MAD contained a larger percentage of significant results. For all three data sets no tendencies were found when comparing the asymmetry indices and kurtoses of treatment C and D. However, there was observed a tendency when comparing the differences in the asymmetry indices and kurtoses of the three data sets. For the data set containing all data points it was found that the data was strongly right skewed-leptokurtic for nearly all behaviors and observation sessions for both individuals. Contrary to this, it was for the IQR data set observed that the data was right skewed-platykurtic for nearly all behaviors and observations sessions for both individuals. For the MAD data set the data was found to be right skewed for all behaviors and observation sessions for both individuals, whereas the tailedness of the distributions varied among the observation sessions and behaviors. Furthermore, some significant differences were found between the asymmetry indices and kurtoses of the two treatments for the data set containing all data points, whereas only a few significant differences were observed when outliers were removed (table 2).

In the comparison of the two individuals it was found for both treatments that individual 1 had a greater median and variance for the behavior activity on land than that of individual 2 (table 3). The same tendency was observed for treatment C for inactive behavior, whereas for treatment D no tendencies were observed when comparing inactive behavior between the two individuals. When comparing the two individuals' stereotypic behavior for treatment C it was generally found that the median time spent on this behavior was greater for individual 1 than for individual 2. For the same treatment there was a slight tendency showing that the variance of time spent on stereotypic behavior was greater for individual 2 than for individual 1. The differences in the median and variance of time that each individual spent on stereotypic behavior for treatment D were only significant when these variables were greater for individual 2 than for individual 1. When comparing medians and variances of all three data sets, the results mostly showed the same tendencies (table 3). There was a small variation due to a slightly larger amount of significant results for the two data sets where outliers had been removed. No tendencies were found when comparing asymmetry index and kurtosis between the two individuals. However, when comparing these results for all three data sets, it was found that the number of significant differences between the individuals generally decreased when outliers had been removed.

For pooled data an increase in the median time spent on each behavior between treatment C and D could be observed for both individuals and all behaviors except the median time individual 2 spent inside, which showed a decrease between treatment C and D (figure 2). For the three behaviors activity in water, stereotypic and inactive, significant differences in the median time were found between the two treatments for each individual. When comparing the median time spent on stereotypic behavior there was also found a significant difference between the individuals for treatment D but not for treatment C. For the behavior inactive, there was only observed a significant difference between the individuals for treatment C. For most behaviors it was found for both individuals that the variances increased between treatment C and D (figure 2). The opposite was only found for the stereotypic behavior of individual 1 and the behavior inside for individual 2, meaning that the variances decreased between treatment C and D for these combinations. For the behaviors inactive and inside, significant differences were found between the variances of both individuals and between these of the two treatments. There were also found significant differences between the variances of time spent on stereotypic behavior between the two individuals for both treatments and between the two treatments for individual 2.

When comparing the asymmetry indices of the two treatments it varied greatly for both individuals whether there was found an increase or decrease in the asymmetry index between the two treatments (figure 2). Significant differences between the asymmetry indices of the two treatments were found for the behaviors stereotypic, inactive and inside of individual 1, whereas for individual 2 it was only the behavior inside that showed a significant difference. Furthermore, no significant differences were found between the two individuals for either of the two treatments. Similar to that of the asymmetry indices there was also great variation in whether the slope was positive or negative when comparing the kurtoses of the two treatments for the different behaviors (figure 2). For all behaviors it was found that there were significant differences between the kurtoses of the two treatments for both individuals. When comparing the kurtoses of the two individuals for treatment C, significant differences were found for the behaviors activity on land, stereotypic and inactive. For treatment D significant differences were found between the kurtoses of the two individuals for the behaviors activity on land and activity in water.

#### 4. Discussion

In this study, an increase in median time and variance were found for most behaviors when the individuals were exposed to the olfactory stimuli of dog odor. This indicates that the occurrences of a behavior generally lasted longer when the individuals were provided with the olfactory stimuli, but also that the individuals were less predictable in the time they were engaged in each occurrence of a behavior. The effect of stimuli on the asymmetry index and kurtosis varied greatly between the individuals and behaviors. This demonstrates that there was a variation in predictability for the behaviors of both individuals when exposed to the olfactory stimuli. In this study the most notable results were found for inactive and stereotypic behavior, and since stereotypy is described as an important indicator of welfare the discussion of the results found in this study will focus on these two behaviors.

This study found that one individual spent significantly more time being inactive under treatment D compared to treatment C, indicating an indirect reaction to the stimulus. In relation to this Ross [11] found, that merely the presence of an activity can influence the behavior of an individual, even if the individual does not engage in the activity. This could explain the influence of stimuli on the individual's inactivity even though this behavior is not directly associated with the presence of the stimulus. Carlstead and Seidensticker [12] also used olfactory stimuli and showed that the presence of these stimuli decreased the time black bears used on stereotypic behavior. The same tendency was observed in this study for one of the individuals (figure 1). The negative correlation between enrichment and time spent on stereotypic behavior has also been shown in various other studies [9, 13]. Contrary to this, the time spent by the other individual on stereotypic behavior increased in the presence of the olfactory stimuli. Shyne [13] found that multiple other studies have shown that stereotypic behavior decreased in the presence of enrichment. This effect could be explained by stereotypy being a general coping mechanism, meaning that stereotypy is not necessarily an indication of low welfare but simply an individual's way of coping. Therefore, stereotypy may help the individual deal with environmental stress and improve the individual's welfare short term. The type of stereotypy used to cope with environmental stress has previously been shown to be species-specific [7, 8]. However, the only way to determine whether stereotypy has a negative or positive effect on welfare is to conduct physiological assessments, such as measuring corticoid levels. Shepherdson et al. [9] studied the correlation between stereotypy and corticoid levels in polar bears and found that pacing generally resulted in higher corticoid levels, which are associated with diminished welfare.

The difference found in the two individuals' responses to olfactory stimuli is a good example of how individuals can respond differently to environmental stress indicating a difference in their personalities. This exhibits how the understanding of different personalities is important in the evaluation of welfare in captive animals [5], implying that different individuals can benefit from different types of enrichment in order to increase their welfare. When exposed to olfactory stimuli, there was a significant difference between the two individuals in the amount of time each spent on stereotypic and inactive behavior, respectively (Appendix B). One individual spent less time being stereotypic and more time on inactivity, while the other individual spent less time being inactive and more time on stereotypy (figure 1). The same was found when comparing the quantitative variables; median, variance, asymmetry index and kurtosis, of the data for the two individuals. This analysis showed differences in medians and variances of time spent on stereotypic behavior between the individuals and that these differences were larger when the individuals were exposed to olfactory stimuli (figure 2). This demonstrates that the individuals responded differently to the stimulus, supporting the statement that individuals with different personalities react differently to the same stimulus, as they often have different ways of coping with changes in their environment [7]. When comparing the asymmetry indices of both individuals it was found that there was a smaller difference between the individuals when exposed to stimuli than under normal conditions, this means that the distributions were more similar. These various results indicate the importance of using different quantitative variables.

The results discussed were generally observed for all three data sets, but some slight differences were found due to the removal of outliers. When using the MAD method only large values were indicated as outliers and removed due to the distribution of the data, whereas when using the IQR to identify outliers an equal amount of values smaller and larger than the median was removed. When removing outliers using IQR, only the most frequent results are shown, and it can be argued that this gives a better representation of the data. A similar argument presents itself when removing outliers using MAD, as this method removes extreme values that have a small likelihood of

occurring. When studying behavior, the distribution of the data is usually skewed to the right, hence the removal of outliers using MAD can remove important information, as an individual performing a behavior for a long time is also a part of their behavior and cannot simply be ignored [24]. Even though removing outliers presents some disadvantages it can also be a resourceful tool when comparing individuals and treatments since the removal of outliers can increase the amount of significant results. The results of this study showed that when removing outliers, tendencies that were found for the data set containing all data points were highlighted with a greater amount of significant differences between the comparisons of the medians and variances of the two treatments and that of the two individuals. This is an example of how removing outliers can be used as a resourceful tool, but the original data set should always be analyzed as well. Ideally, behavioral data should be analyzed both with and without outliers, as the different methods supplement each other.

In this study it is shown how behavioral instability can be applied to behavioral observations by investigating the medians, variances, asymmetry indices and kurtoses of different behaviors. The results of this study prove that behavioral instability also can be applied to a more traditional type of behavioral data, in which an ethogram and observations of different behaviors are used. Thus, behavioral instability can be introduced as a new quantitative method for analyzing traditional ethograms. This study used this new method along with the traditional methods, enabling a comparison of the two methods. One of the major issues when using the traditional methods for studying behavior is the lack of comparable systematic and quantitative results [8, 13]. The traditional methods are primarily used to estimate the percent of time spent on various activities [15]. This estimate is however highly dependent on the ethogram used, as the percent of time spent on one activity is always dependent on the amount of time spent on other activities. Comparisons between studies are therefore only possible if highly similar ethograms are used, which can prove difficult in the comparison of behavior between species. The application of the concept behavioral instability enables the comparison of behavior regardless of differences in ethograms. This is possible due to the method's quantitative approach that uses the median, variance, asymmetry index and kurtosis. The advantage of this approach is that these variables for one behavior are less dependent on the other behaviors.

The traditional methods also lack a protocol ensuring systematic data sampling. The results of this study indicate the need for longer observation sessions, as short observation sessions lead to a higher risk of type II errors. However, if the data is symmetric-leptokurtic the risk of type II errors is lower, thus it is possible to make a behavioral analysis based on short observation sessions. When applying the concept of behavioral instability to behavioral studies, the data should be sampled using continuous focal sampling of the entire day. The results of this study showed that many behaviors occurred for both shorter and longer periods of time and therefore, information can be lost when using sampling techniques such as instantaneous sampling. Altmann [15] states that instantaneous sampling is primarily used for studying the percent of time spent on various activities. However, the results would not be accurate, as behaviors shorter than the time between two preselected sampling instances would most likely not be recorded. When using this new quantitative method of applying behavioral instability, it is therefore important that sampling is conducted throughout the entire day using continuous focal sampling.

While the quantitative results of this new method enable comparisons between studies, the traditional methods should not be dismissed, as valuable information also lies in knowing when an individual performs various behaviors throughout the day and the proportion of the day spent on different behaviors. It is therefore suggested that the two methods are used collaboratively, comparing and combining the results of both approaches, in order to obtain the most reliable results. The application of the concept behavioral instability to traditional behavioral analyses allows quantitative data collection. This can provide researchers with a relatively unbiased evaluation of personality and the effectiveness of enrichment manipulation, which can contribute to the improvement of enrichment programs and animal welfare in captivity [13]. It has been debated whether the study of personality can provide new insights for the field of behavioral ecology [5]. The use of behavioral instability as a new quantitative and systematic method for studying personality could be highly relevant when studying animal conservation. When captive populations are being managed with the purpose of re-introducing individuals to the wild, the understanding of personality can provide insight on how to conserve behaviors and personality traits that could be beneficial in the wild.

**Acknowledgments:** The authors wish to thank the staff of Aalborg Zoo, with special thanks to Katrine Christensen and Frank Thomsen for their time and involvement in this study.

## REFERENCES

1. Patrick Myers and Julie Young. Consistent Individual Behavior: Evidence of Personality in Black Bears. *Journal of Ethology*, 36(2):117–124, 2018.
2. Francesca Santicchia, Lucas A. Wauters, Ben Dantzer, Sarah E. Westrick, Nicola Ferrari, Claudia Romeo, Rupert Palme, Damiano G. Preatoni and Adriano Martinoli. Relationships Between Personality Traits and the Physiological Stress Response in a Wild Mammal. *Current Zoology*, 0(0):1–8, 2019.
3. Pierre-Olivier Montiglio, Dany Garant, Fanie Pelletier, and Denis Réale. Personality Differences are Related to Long-Term Stress Reactivity in a Population of Wild Eastern Chipmunks, *Tamias striatus*. *Animal Behaviour*, 84(4):1071–1079, 2012.
4. Lu Luo, Inonge Reimert, Elske N. de Haas, Bas Kemp, and J. E. Bolhuis. Effects of Early and Later Life Environmental Enrichment and Personality on Attention Bias in Pigs (*Sus scrofa domesticus*). *Animal Cognition*, 22(6):959–972, 2019.
5. Vanessa Wilson, Anja Guenther, Øyvind Øverli, Martin W. Seltmann, and Drew Altschul. Future Directions for Personality Research: Contributing New Insights to the Understanding of Animal Behavior. *Animals*, 9 (240):2–17, 2019.
6. Ros Clubb and Georgia J. Mason. Animal Welfare: Captivity Effects on Wide-Ranging Carnivores. *Nature*, 425(6957):473–474, 2003.
7. Georgia J. Mason and Naomi R. Latham. Can't Stop, Won't Stop: Is Stereotypy a Reliable Animal Welfare Indicator? *Animal Welfare*, 13(1):57–69, 2004.
8. Paul E. Rose, Steve M. Nash, and Lisa M. Riley. To Pace or Not to Pace? A Review of What Abnormal Repetitive Behavior Tells Us About Zoo Animal Management. *Journal of Veterinary Behavior*, 20:11–21, 2017.
9. David Shepherdson, Karen D. Lewis, Kathy Carlstead, Joan Bauman, and Nancy Perrin. Individual and Environmental Factors Associated with Stereotypic Behavior and Fecal Glucocorticoid Metabolite Levels in Zoo Housed Polar Bears. *Applied Animal Behavior Science*, 147(3-4):268–277, 2013.
10. Isabelle T. Cless and Kristen E. Lukas. Variables Affecting the Manifestation of and Intensity of Pacing Behavior: A Preliminary Case Study in Zoo-Housed Polar Bears. *Zoo Biology*, 36(5):307–315, 2017.
11. Stephen R. Ross. Issues of Choice and Control in the Behaviour of a Pair of Captive Polar Bears *Ursus maritimus*. *Behavioral Processes*, 73(1):117–120, 2006.
12. Kathy Carlstead and John Seidensticker. Seasonal Variation in Stereotypic Pacing in an American Black Bear *Ursus americanus*. *Behavioural Processes*, 25:155–161, 1991.
13. Amanda Shyne. Meta-Analytic Review of the Effects of Enrichment on Stereotypic Behavior in Zoo Mammals. *Zoo Biology*, 25(4):317–337, 2006.
14. Isabelle T. Cless, Heather A. Voss-Hoynes, Roy E. Ritzmann, and Kristen E. Lukas. Defining Pacing Quantitatively: A Comparison of Gait Characteristics Between Pacing and Non-repetitive Locomotion in Zoo-Housed Polar Bears. *Applied Animal Behaviour Science*, 169:78–85, 2015.
15. Jeanne Altmann. Observational Study of Behavior: Sampling Methods. *Behaviour*, 49(3/4):227–267, 1974.
16. Alison Ames. Object Manipulation in Captive Polar Bears. *Int. Conf. Bear Res. and Manage*, 9(1):443–449, 1994.
17. Meredith J. Bashaw, Angela S. Kelling, Mollie A. Bloomsmith, and Terry L. Maple. Environmental Effects on the Behavior of Zoo-Housed Lions and Tigers, with a Case Study of the Effects of a Visual Barrier on Pacing. *Journal of Applied Animal Welfare Science*, 10(2):95–109, 2007.
18. Cino Pertoldi, Simon Bahrndorff, Zorana K. Novicic, and Palle D. Rohde. The Novel Concept of "Behavioral Instability" and its Potential Applications. *Symmetry*, 8(135):1–13, 2016.
19. Mads Bech-Hansen, Rune M. Kallehauge, Dan Bruhn, Johan H. F. Castenschiold, Jonas B. Gehrlein, Bjarke Laubek, Lasse F. Jensen, and Cino Pertoldi. Effect of Landscape Elements on the Symmetry and Variance of the Spatial Distribution of Individual Birds within Foraging Flocks of Geese. *Symmetry*, 11(1103):1–11, 2019.
20. RStudio Team. RStudio: Integrated Development for R. *RStudio, Inc.*, 2016.
21. Øyvind Hammer, David A. T. Harper, and Paul D. Ryan. Past: Paleontological Statistics Software Package for Education and Data Analysis. *Palaeontologia Electronica*, 4(1), 2001.
22. Christophe Leys, Christophe Ley, Olivier Klein, Philippe Bernard, and Laurent Licata. Detecting Outliers: Do Not Use Standard Deviation around the Mean, Use Absolute Deviation around the Median. *Journal of Experimental Social Psychology*, 49(4):764–766, 2013.
23. Jerrold H. Zar. *Biostatistical Analysis*. Pearson, Northern Illinois University, 4<sup>th</sup> edition, 1999.



24. Peter J. Rousseeuw and Christophe Croux. Alternatives to the median absolute deviation. *Journal of the American Statistical Association*, 88(424):1273–1283, 1993.

386

**Table 1.** Behavioral ethogram

387

388

389

390

391

392

393

394

395

396

397

398

399

<i><b>Behavior</b></i>	<i><b>Description</b></i>
Activity on Land	Locomotion and interaction with objects while on land
Activity in Water	Locomotion and interaction with objects while submerged in water
Social Play	Individuals interacting playfully or fighting with each other
Stereotypic	Repeating a specific walking pattern or movement aimlessly
Inactive	Resting or sleeping i.e. laying down or sitting with minimal movement
Inside	Inside the den and therefore out of sight
Other	Eating, drinking, urinating, defecating, maintenance of coat e.g. by rolling in gravel and out of sight due to blind camera angles

**Table 2.** Pairwise comparisons of the median, variance, asymmetry index and kurtosis of each observation session between the two treatments (C = control, D = dog scented object) for each individual. For each behavior results are shown for three data sets, the first including all data points (All), the second only including data points within the interquartile range (IQR) and the third with outliers removed using the median absolute deviation method (MAD). The shown percentage indicates the proportion of significant results for the given difference. The note column indicates the number of observation sessions that have less than 6 data points for the given behavior.

Behavior	Individual	Median			Variance			Skewness			Kurtosis			Note
		All	IQR	MAD	All	IQR	MAD	All	IQR	MAD	All	IQR	MAD	
Activity on Land	1	C > D: 0%	C > D: 0%	C > D: 0%	C > D: 0%	C > D: 0%	C > D: 0%	C > D: 11%	C > D: 0%	C > D: 0%	C > D: 11%	C > D: 0%	C > D: 0%	
		C < D: 56%	C < D: 67%	C < D: 67%	C < D: 33%	C < D: 67%	C < D: 67%	C < D: 0%	C < D: 0%	C < D: 0%	C < D: 0%	C < D: 0%	C < D: 0%	
	2	C > D: 0%	C > D: 22%	C > D: 22%	C > D: 0%	C > D: 22%	C > D: 33%	C > D: 11%	C > D: 0%	C > D: 0%	C > D: 11%	C > D: 0%	C > D: 0%	
		C < D: 33%	C < D: 44%	C < D: 33%	C < D: 44%	C < D: 44%	C < D: 44%	C < D: 0%	C < D: 0%	C < D: 0%	C < D: 0%	C < D: 0%	C < D: 0%	
Activity in Water	1	C > D: 0%	C > D: 0%	C > D: 0%	C > D: 0%	C > D: 0%	C > D: 11%	C > D: 0%	C > D: 0%	C > D: 0%	C > D: 11%	C > D: 0%	C > D: 0%	
		C < D: 22%	C < D: 33%	C < D: 33%	C < D: 44%	C < D: 44%	C < D: 22%	C < D: 0%	C < D: 0%	C < D: 0%	C < D: 0%	C < D: 0%	C < D: 0%	
	2	C > D: 0%	C > D: 11%	C > D: 11%	C > D: 0%	C > D: 0%	C > D: 11%	C > D: 0%	C > D: 0%	C > D: 0%	C > D: 0%	C > D: 0%	C > D: 0%	IQR: 1D
		C < D: 33%	C < D: 78%	C < D: 44%	C < D: 22%	C < D: 33%	C < D: 44%	C < D: 11%	C < D: 0%	C < D: 0%	C < D: 22%	C < D: 0%	C < D: 0%	MAD: 1D
Social Play	Both	C > D: 33%	C > D: 67%	C > D: 33%	C > D: 22%	C > D: 44%	C > D: 56%	C > D: 0%	C > D: 0%	C > D: 0%	C > D: 0%	C > D: 0%	C > D: 0%	
		C < D: 0%	C < D: 0%	C < D: 11%	C < D: 0%	C < D: 0%	C < D: 11%	C < D: 44%	C < D: 0%	C < D: 0%	C < D: 44%	C < D: 0%	C < D: 0%	
Stereotypic	1	C > D: 22%	C > D: 22%	C > D: 22%	C > D: 0%	C > D: 11%	C > D: 22%	C > D: 0%	C > D: 0%	C > D: 11%	C > D: 0%	C > D: 0%	C > D: 0%	
		C < D: 22%	C < D: 33%	C < D: 33%	C < D: 11%	C < D: 22%	C < D: 33%	C < D: 0%	C < D: 0%	C < D: 0%	C < D: 0%	C < D: 0%	C < D: 0%	
	2	C > D: 0%	C > D: 0%	C > D: 0%	C > D: 0%	C > D: 0%	C > D: 11%	C > D: 0%	C > D: 0%	C > D: 0%	C > D: 0%	C > D: 0%	C > D: 0%	IQR: 1D
		C < D: 11%	C < D: 67%	C < D: 44%	C < D: 22%	C < D: 44%	C < D: 55%	C < D: 0%	C < D: 22%	C < D: 0%	C < D: 0%	C < D: 22%	C < D: 0%	MAD: 1D

---

<b>Inactive</b>	<b>1</b>	C > D: 0%	C > D: 0%	C > D: 0%	C > D: 0%	C > D: 0%	C > D: 0%	C > D: 0%	C > D: 0%	C > D: 0%	C > D: 0%	C > D: 0%	C > D: 0%	All: 1C, 1D
		C < D: 11%	C < D: 33%	C < D: 33%	C < D: 22%	C < D: 33%	C < D: 33%	C < D: 11%	C < D: 0%	C < D: 0%	C < D: 33%	C < D: 0%	C < D: 0%	IQR: 1C, 1D
	<b>2</b>	C > D: 0%	C > D: 0%	C > D: 0%	C > D: 0%	C > D: 0%	C > D: 0%	C > D: 0%	C > D: 0%	C > D: 0%	C > D: 0%	C > D: 0%	C > D: 0%	All: 1C, 2D
		C < D: 22%	C < D: 22%	C < D: 22%	C < D: 11%	C < D: 22%	C < D: 44%	C < D: 11%	C < D: 0%	C < D: 0%	C < D: 22%	C < D: 0%	C < D: 0%	IQR: 2C, 2D
														MAD: 1C, 2D

---

<b>Inside</b>	<b>1</b>	C > D: 0%	C > D: 0%	C > D: 0%	C > D: 0%	C > D: 0%	C > D: 11%	C > D: 0%	C > D: 0%	C > D: 33%	C > D: 0%	C > D: 0%	All: 1C	
		C < D: 22%	C < D: 33%	C < D: 33%	C < D: 0%	C < D: 0%	C < D: 22%	C < D: 0%	C < D: 0%	C < D: 0%	C < D: 11%	C < D: 0%	C < D: 11%	IQR: 1C, 2D
	<b>2</b>	C > D: 11%	C > D: 33%	C > D: 11%	C > D: 22%	C > D: 22%	C > D: 44%	C > D: 33%	C > D: 0%	C > D: 0%	C > D: 44%	C > D: 0%	C > D: 0%	IQR: 1D
		C < D: 33%	C < D: 0%	C < D: 33%	C < D: 0%	C < D: 0%	C < D: 33%	C < D: 11%	C < D: 0%	C < D: 0%	C < D: 11%	C < D: 0%	C < D: 0%	

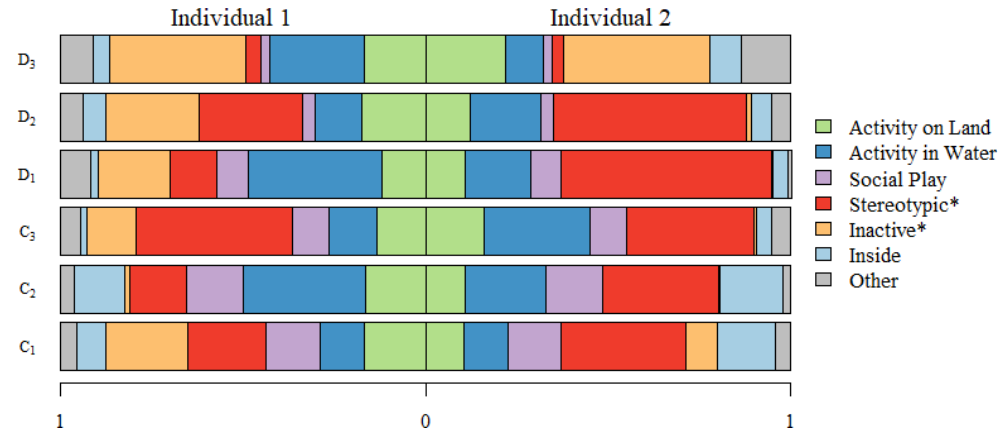
---

**Table 3.** Pairwise comparisons of the median, variance, asymmetry index and kurtoses of each observation session between the two individuals ( $I_1$  = individual 1,  $I_2$  = individual 2) for each treatment. For each behavior results are shown for three data sets, the first including all data points (All), the second only including data points within the interquartile range (IQR) and the third where outliers were removed using the median absolute deviation method (MAD). The shown percentage indicates the proportion of significant results for the given difference. The note column indicates the number of observation sessions that have less than 6 data points for the given behavior.

Behavior	Treatment	Median			Variance			Skewness			Kurtosis			Note		
		All	MAD	IQR	All	MAD	IQR	All	MAD	IQR	All	MAD	IQR			
Activity on Land	C	$I_1 > I_2$ : 67%	$I_1 > I_2$ : 67%	$I_1 > I_2$ : 67%	$I_1 > I_2$ : 22%	$I_1 > I_2$ : 67%	$I_1 > I_2$ : 67%	$I_1 > I_2$ : 0%	$I_1 > I_2$ : 0%	$I_1 > I_2$ : 0%	$I_1 > I_2$ : 0%	$I_1 > I_2$ : 0%	$I_1 > I_2$ : 0%	$I_1 > I_2$ : 0%		
		$I_1 < I_2$ : 0%	$I_1 < I_2$ : 0%	$I_1 < I_2$ : 0%	$I_1 < I_2$ : 0%	$I_1 < I_2$ : 22%	$I_1 < I_2$ : 33%	$I_1 < I_2$ : 0%	$I_1 < I_2$ : 0%	$I_1 < I_2$ : 0%	$I_1 < I_2$ : 0%	$I_1 < I_2$ : 0%	$I_1 < I_2$ : 0%	$I_1 < I_2$ : 0%		
	D	$I_1 > I_2$ : 56%	$I_1 > I_2$ : 56%	$I_1 > I_2$ : 67%	$I_1 > I_2$ : 22%	$I_1 > I_2$ : 67%	$I_1 > I_2$ : 56%	$I_1 > I_2$ : 0%	$I_1 > I_2$ : 0%	$I_1 > I_2$ : 0%	$I_1 > I_2$ : 0%	$I_1 > I_2$ : 0%	$I_1 > I_2$ : 0%	$I_1 > I_2$ : 0%		
		$I_1 < I_2$ : 0%	$I_1 < I_2$ : 0%	$I_1 < I_2$ : 11%	$I_1 < I_2$ : 11%	$I_1 < I_2$ : 11%	$I_1 < I_2$ : 22%	$I_1 < I_2$ : 11%	$I_1 < I_2$ : 0%	$I_1 < I_2$ : 0%	$I_1 < I_2$ : 0%	$I_1 < I_2$ : 11%	$I_1 < I_2$ : 0%	$I_1 < I_2$ : 0%		
	Activity in Water	C	$I_1 > I_2$ : 22%	$I_1 > I_2$ : 44%	$I_1 > I_2$ : 33%	$I_1 > I_2$ : 11%	$I_1 > I_2$ : 33%	$I_1 > I_2$ : 11%	$I_1 > I_2$ : 0%	$I_1 > I_2$ : 0%	$I_1 > I_2$ : 0%	$I_1 > I_2$ : 0%	$I_1 > I_2$ : 0%	$I_1 > I_2$ : 0%	$I_1 > I_2$ : 0%	
			$I_1 < I_2$ : 0%	$I_1 < I_2$ : 22%	$I_1 < I_2$ : 22%	$I_1 < I_2$ : 0%	$I_1 < I_2$ : 22%	$I_1 < I_2$ : 22%	$I_1 < I_2$ : 0%	$I_1 < I_2$ : 0%	$I_1 < I_2$ : 0%	$I_1 < I_2$ : 0%	$I_1 < I_2$ : 0%	$I_1 < I_2$ : 0%	$I_1 < I_2$ : 0%	
D		$I_1 > I_2$ : 22%	$I_1 > I_2$ : 22%	$I_1 > I_2$ : 22%	$I_1 > I_2$ : 0%	$I_1 > I_2$ : 22%	$I_1 > I_2$ : 22%	$I_1 > I_2$ : 0%	$I_1 > I_2$ : 0%	$I_1 > I_2$ : 0%	$I_1 > I_2$ : 0%	$I_1 > I_2$ : 0%	$I_1 > I_2$ : 0%	$I_1 > I_2$ : 0%	IQR: 1 $I_2$	
		$I_1 < I_2$ : 11%	$I_1 < I_2$ : 22%	$I_1 < I_2$ : 0%	$I_1 < I_2$ : 0%	$I_1 < I_2$ : 11%	$I_1 < I_2$ : 11%	$I_1 < I_2$ : 11%	$I_1 < I_2$ : 0%	$I_1 < I_2$ : 0%	$I_1 < I_2$ : 0%	$I_1 < I_2$ : 33%	$I_1 < I_2$ : 0%	$I_1 < I_2$ : 0%	MAD: 1 $I_2$	
Stereotypic	C	$I_1 > I_2$ : 22%	$I_1 > I_2$ : 33%	$I_1 > I_2$ : 33%	$I_1 > I_2$ : 0%	$I_1 > I_2$ : 22%	$I_1 > I_2$ : 11%	$I_1 > I_2$ : 0%	$I_1 > I_2$ : 0%	$I_1 > I_2$ : 0%	$I_1 > I_2$ : 0%	$I_1 > I_2$ : 0%	$I_1 > I_2$ : 0%	$I_1 > I_2$ : 0%		
		$I_1 < I_2$ : 11%	$I_1 < I_2$ : 0%	$I_1 < I_2$ : 11%	$I_1 < I_2$ : 33%	$I_1 < I_2$ : 22%	$I_1 < I_2$ : 22%	$I_1 < I_2$ : 0%	$I_1 < I_2$ : 0%	$I_1 < I_2$ : 0%	$I_1 < I_2$ : 0%	$I_1 < I_2$ : 0%	$I_1 < I_2$ : 0%	$I_1 < I_2$ : 0%		
	D	$I_1 > I_2$ : 0%	$I_1 > I_2$ : 0%	$I_1 > I_2$ : 0%	$I_1 > I_2$ : 0%	$I_1 > I_2$ : 0%	$I_1 > I_2$ : 0%	$I_1 > I_2$ : 0%	$I_1 > I_2$ : 0%	$I_1 > I_2$ : 0%	$I_1 > I_2$ : 33%	$I_1 > I_2$ : 0%	$I_1 > I_2$ : 0%	$I_1 > I_2$ : 22%	MAD: 1 $I_2$	
		$I_1 < I_2$ : 11%	$I_1 < I_2$ : 11%	$I_1 < I_2$ : 44%	$I_1 < I_2$ : 44%	$I_1 < I_2$ : 67%	$I_1 < I_2$ : 44%	$I_1 < I_2$ : 0%	$I_1 < I_2$ : 11%	$I_1 < I_2$ : 0%	$I_1 < I_2$ : 0%	$I_1 < I_2$ : 11%	$I_1 < I_2$ : 0%	$I_1 < I_2$ : 0%	IQR: 1 $I_2$	
Inactive	C	$I_1 > I_2$ : 11%	$I_1 > I_2$ : 22%	$I_1 > I_2$ : 33%	$I_1 > I_2$ : 11%	$I_1 > I_2$ : 44%	$I_1 > I_2$ : 22%	$I_1 > I_2$ : 11%	$I_1 > I_2$ : 0%	$I_1 > I_2$ : 0%	$I_1 > I_2$ : 22%	$I_1 > I_2$ : 0%	$I_1 > I_2$ : 0%	All: 1 $I_1$ ,		
		$I_1 < I_2$ : 0%	$I_1 < I_2$ : 0%	$I_1 < I_2$ : 0%	$I_1 < I_2$ : 0%	$I_1 < I_2$ : 0%	$I_1 < I_2$ : 0%	$I_1 < I_2$ : 0%	$I_1 < I_2$ : 0%	$I_1 < I_2$ : 0%	$I_1 < I_2$ : 0%	$I_1 < I_2$ : 0%	$I_1 < I_2$ : 0%	$I_1 < I_2$ : 0%	1 $I_2$	

---

																		MAD: 1I <sub>1</sub>
																		IQR: 1I <sub>1</sub> ,
																		2I <sub>2</sub>
<b>D</b>	I <sub>1</sub> > I <sub>2</sub> : 11%	I <sub>1</sub> > I <sub>2</sub> : 11%	I <sub>1</sub> > I <sub>2</sub> : 11%	I <sub>1</sub> > I <sub>2</sub> : 0%	I <sub>1</sub> > I <sub>2</sub> : 0%	I <sub>1</sub> > I <sub>2</sub> : 0%	I <sub>1</sub> > I <sub>2</sub> : 11%	I <sub>1</sub> > I <sub>2</sub> : 0%	I <sub>1</sub> > I <sub>2</sub> : 0%	I <sub>1</sub> > I <sub>2</sub> : 67%	I <sub>1</sub> > I <sub>2</sub> : 0%	I <sub>1</sub> > I <sub>2</sub> : 0%	All: 1I <sub>1</sub> ,					
	I <sub>1</sub> < I <sub>2</sub> : 0%	I <sub>1</sub> < I <sub>2</sub> : 11%	I <sub>1</sub> < I <sub>2</sub> : 0%	I <sub>1</sub> < I <sub>2</sub> : 0%	I <sub>1</sub> < I <sub>2</sub> : 11%	I <sub>1</sub> < I <sub>2</sub> : 0%	I <sub>1</sub> < I <sub>2</sub> : 0%	I <sub>1</sub> < I <sub>2</sub> : 0%	I <sub>1</sub> < I <sub>2</sub> : 0%	I <sub>1</sub> < I <sub>2</sub> : 22%	I <sub>1</sub> < I <sub>2</sub> : 0%	I <sub>1</sub> < I <sub>2</sub> : 0%	2I <sub>2</sub>					
													MAD: 2I <sub>2</sub>					
													IQR: 1I <sub>1</sub> ,					
													2I <sub>2</sub>					
<b>C</b>	I <sub>1</sub> > I <sub>2</sub> : 0%	I <sub>1</sub> > I <sub>2</sub> : 0%	I <sub>1</sub> > I <sub>2</sub> : 0%	I <sub>1</sub> > I <sub>2</sub> : 0%	I <sub>1</sub> > I <sub>2</sub> : 0%	I <sub>1</sub> > I <sub>2</sub> : 0%	I <sub>1</sub> > I <sub>2</sub> : 22%	I <sub>1</sub> > I <sub>2</sub> : 0%	I <sub>1</sub> > I <sub>2</sub> : 0%	I <sub>1</sub> > I <sub>2</sub> : 22%	I <sub>1</sub> > I <sub>2</sub> : 0%	I <sub>1</sub> > I <sub>2</sub> : 11%	All: 1I <sub>1</sub>					
<b>Inside</b>	I <sub>1</sub> < I <sub>2</sub> : 22%	I <sub>1</sub> < I <sub>2</sub> : 11%	I <sub>1</sub> < I <sub>2</sub> : 33%	I <sub>1</sub> < I <sub>2</sub> : 0%	I <sub>1</sub> < I <sub>2</sub> : 44%	I <sub>1</sub> < I <sub>2</sub> : 44%	I <sub>1</sub> < I <sub>2</sub> : 11%	I <sub>1</sub> < I <sub>2</sub> : 0%	I <sub>1</sub> < I <sub>2</sub> : 0%	I <sub>1</sub> < I <sub>2</sub> : 11%	I <sub>1</sub> < I <sub>2</sub> : 0%	I <sub>1</sub> < I <sub>2</sub> : 0%	IQR: 1I <sub>1</sub>					
<b>D</b>	I <sub>1</sub> > I <sub>2</sub> : 11%	I <sub>1</sub> > I <sub>2</sub> : 22%	I <sub>1</sub> > I <sub>2</sub> : 22%	I <sub>1</sub> > I <sub>2</sub> : 11%	I <sub>1</sub> > I <sub>2</sub> : 11%	I <sub>1</sub> > I <sub>2</sub> : 11%	I <sub>1</sub> > I <sub>2</sub> : 11%	I <sub>1</sub> > I <sub>2</sub> : 0%	I <sub>1</sub> > I <sub>2</sub> : 0%	I <sub>1</sub> > I <sub>2</sub> : 33%	I <sub>1</sub> > I <sub>2</sub> : 0%	I <sub>1</sub> > I <sub>2</sub> : 11%	IQR: 2I <sub>1</sub> ,					
	I <sub>1</sub> < I <sub>2</sub> : 22%	I <sub>1</sub> < I <sub>2</sub> : 22%	I <sub>1</sub> < I <sub>2</sub> : 22%	I <sub>1</sub> < I <sub>2</sub> : 11%	I <sub>1</sub> < I <sub>2</sub> : 33%	I <sub>1</sub> < I <sub>2</sub> : 22%	I <sub>1</sub> < I <sub>2</sub> : 0%	I <sub>1</sub> < I <sub>2</sub> : 0%	I <sub>1</sub> < I <sub>2</sub> : 0%	I <sub>1</sub> < I <sub>2</sub> : 0%	I <sub>1</sub> < I <sub>2</sub> : 22%	I <sub>1</sub> < I <sub>2</sub> : 0%	1I <sub>2</sub>					

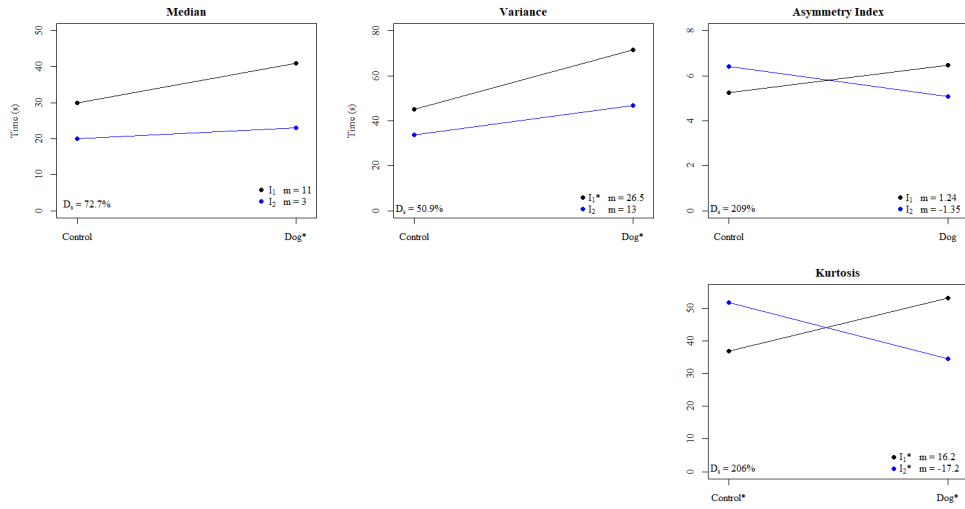


**Figure 1.** Ratio of time each individual spent on the different behaviors for each of the three observation sessions for each treatment (C = control, D = dog scented object). The data was pooled and compared by  $\chi^2$  tests with Yates correction. Behaviors in which the  $\chi^2$  test resulted in significant results are indicated by \*. For further details on  $\chi^2$  values see Appendix B.

1  
2

a) Activity on Land

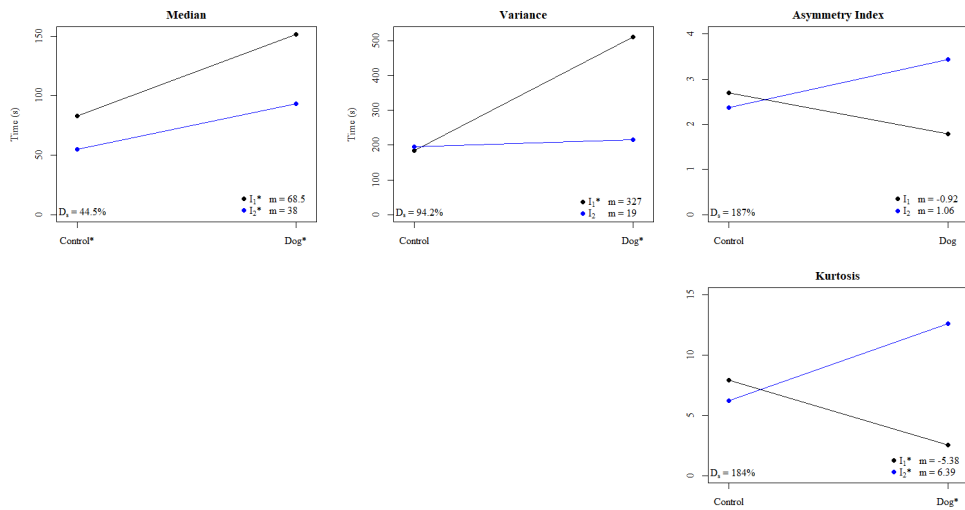
3



4  
5  
6  
7

b) Activity in Water

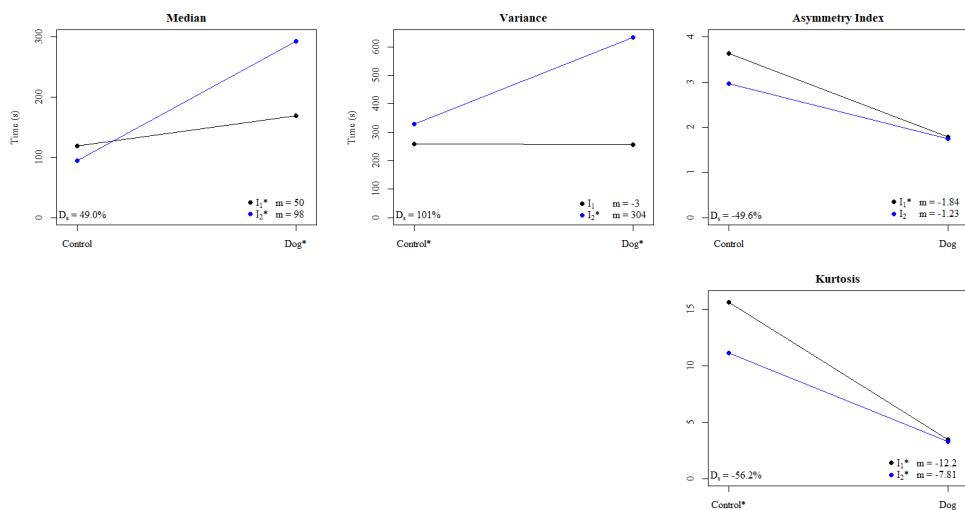
8



9  
10  
11  
12

c) Stereotypic

13

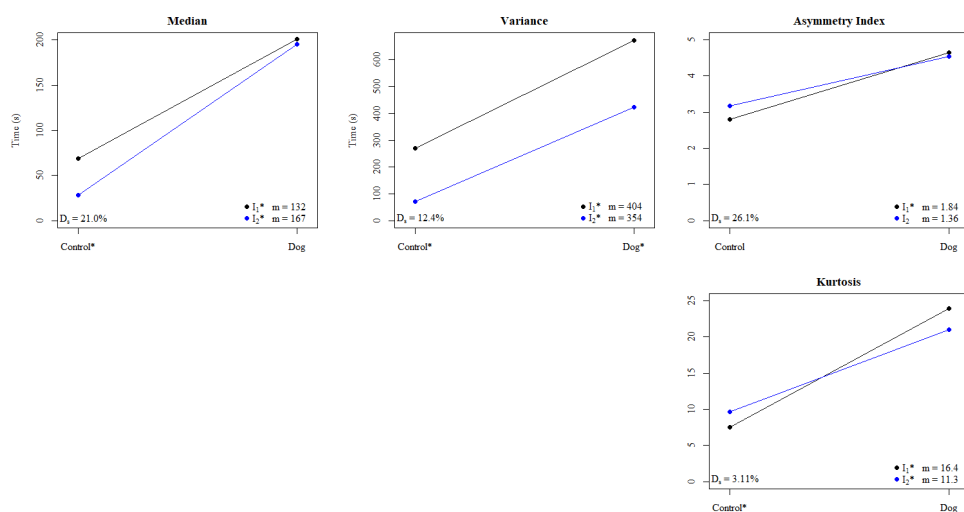


14



15  
16  
17

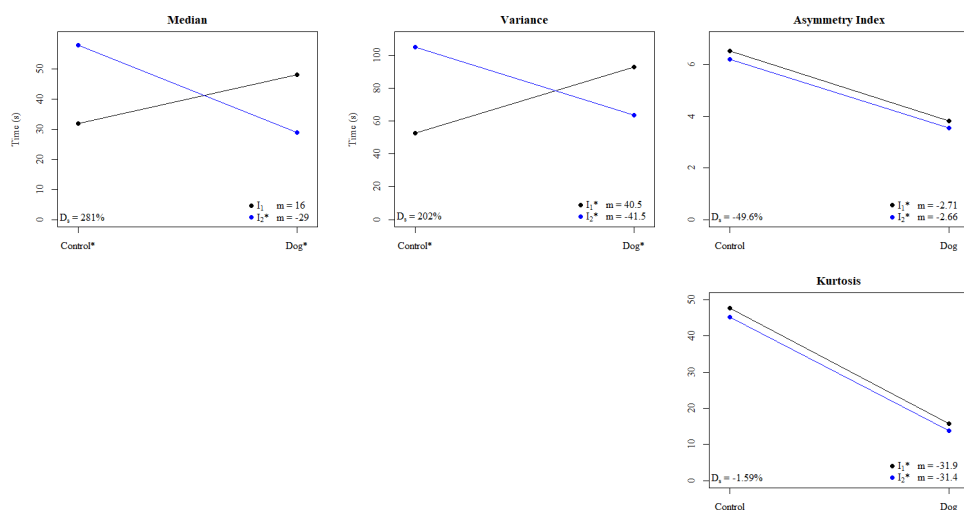
d) Inactive



18

19  
20  
21  
22

e) Inside



23

24  
25

26 **Figure 2.** For each individual the median and variance of time spent on a given behavior are shown for  
 27 treatment C and for treatment D along with trend lines between the medians and between the variances of the  
 28 two treatments for the same individual. The asymmetry index and kurtosis are also shown for treatment C and  
 29 for treatment D and each individual along with trend lines between the asymmetry indices and between the  
 30 kurtoses of the two treatments for the same individual. The medians, variances, asymmetry indices and kurtoses  
 31 are based on pooled data. The slope ( $m$ ) and difference in slope in percent ( $D_s$ ) are given for each comparison.  
 32 The medians, variances, asymmetry indices and kurtoses were compared by  $\chi^2$  tests with Yates correction.  
 33 Comparisons in which the  $\chi^2$  test resulted in significant results between the two individuals for the same  
 34 treatment are indicated by \* next to the relative treatment. Comparisons in which the  $\chi^2$  test resulted in  
 35 significant results between the two treatments for the same individual are indicated by \* next to the relative  
 36 individual. For further details on  $\chi^2$  values see Appendix D.

37