



This is a peer-reviewed, post-print (final draft post-refereeing) version of the following published document:

Stafford, Richard, Goodenough, Anne E ORCID logoORCID: <https://orcid.org/0000-0002-7662-6670>, Slater, Kathy, Carpenter, William S ORCID logoORCID: <https://orcid.org/0009-0001-9031-5561>, Collins, Laura, Cruickshank, Heather, Downing, Sarah, Hall, Sally, McDonald, Katie, McDonnell, Heather, Overs, Bryony, Spiers, Lizzie, Stetter, Mark, Zitzer, Heike and Hart, Adam G ORCID logoORCID: <https://orcid.org/0000-0002-4795-9986> (2011) Inferential and visual analysis of ethogram data using multivariate techniques. *Animal Behaviour*, 83 (2). pp. 563-569. doi:10.1016/j.anbehav.2011.11.020

Official URL: <http://dx.doi.org/10.1016/j.anbehav.2011.11.020>

DOI: <http://dx.doi.org/10.1016/j.anbehav.2011.11.020>

EPrint URI: <https://eprints.glos.ac.uk/id/eprint/3363>

Disclaimer

The University of Gloucestershire has obtained warranties from all depositors as to their title in the material deposited and as to their right to deposit such material.

The University of Gloucestershire makes no representation or warranties of commercial utility, title, or fitness for a particular purpose or any other warranty, express or implied in respect of any material deposited.

The University of Gloucestershire makes no representation that the use of the materials will not infringe any patent, copyright, trademark or other property or proprietary rights.

The University of Gloucestershire accepts no liability for any infringement of intellectual property rights in any material deposited but will remove such material from public view pending investigation in the event of an allegation of any such infringement.

PLEASE SCROLL DOWN FOR TEXT.

This is a peer-reviewed, post-print (final draft post-refereeing) version of the following published document:

Stafford, Richard and Goodenough, Anne E. and Slater, Kathy and Carpenter, William and Collins, Laura and Cruickshank, Heather and Downing, Sarah and Hall, Sally and McDonald, Katie and McDonnell, Heather and Overs, Bryony and Spiers, Lizzie and Stetter, Mark and Zitzer, Heike and Hart, Adam G. (2011). *Inferential and visual analysis of ethogram data using multivariate techniques*. *Animal Behaviour*, 83 (2), 563-569. ISSN 00033472

Published in *Animal Behaviour*, and available online at:

<http://www.sciencedirect.com/science/article/pii/S0003347211005124>

We recommend you cite the published (post-print) version.

The URL for the published version is <http://dx.doi.org/10.1016/j.anbehav.2011.11.020>

Disclaimer

The University of Gloucestershire has obtained warranties from all depositors as to their title in the material deposited and as to their right to deposit such material.

The University of Gloucestershire makes no representation or warranties of commercial utility, title, or fitness for a particular purpose or any other warranty, express or implied in respect of any material deposited.

The University of Gloucestershire makes no representation that the use of the materials will not infringe any patent, copyright, trademark or other property or proprietary rights.

The University of Gloucestershire accepts no liability for any infringement of intellectual property rights in any material deposited but will remove such material from public view pending investigation in the event of an allegation of any such infringement.

PLEASE SCROLL DOWN FOR TEXT

1 **Inferential and visual analysis of ethogram data using multivariate techniques**

2 Richard Stafford^{a*}, Anne E. Goodenough^b, Kathy Slater^c, William Carpenter^b, Laura Collins^b, Heather
3 Cruickshank^b, Sarah Downing^b, Sally Hall^b, Katie McDonald^d, Heather McDonnell^b, Bryony Overs^b,
4 Lizzie Spiers^b, Mark Stetter^e, Heike Zitzer^c, Adam G. Hart^b

5

6 a) Luton Institute of Research in Applied Natural Sciences, University of Bedfordshire

7 b) Department of Natural and Social Sciences, University of Gloucestershire, Cheltenham

8 c) Operation Wallacea, Spilsby, Lincolnshire

9 d) West Midlands Safari Park, Bewdley, Worcestershire

10 e) Animals Science and Environment, Walt Disney World, Lake Buena Vista, Florida

11

12 * Corresponding author. Email: rick.stafford7@gmail.com

13 Division of Science, University of Bedfordshire, Luton, LU1 3JU, UK.

14 Tel: +44 (0) 1582 489158

15 Fax: +44 (0) 1582 489212

16

17 **Inferential and visual analysis of ethogram data using multivariate techniques**

18 Activity budgets are frequently used to examine behaviours of animals, especially of large mammals
19 in field or captivity conditions (e.g. Altmann 1974; Weller and Bennett 2001; MacNulty et al. 2007).
20 Often, such processes are conducted using ethograms, where a number of typical behaviours are
21 listed (such as foraging, sleeping, walking, standing still, interacting with others) and either the
22 duration of each behaviour within each observation period is noted, or, more normally, the
23 occurrence of a certain behaviour is recorded at a regular time interval (Dawkins 2007; Martin and
24 Bateson 2007). The technique is simple, and clearly effective in calculating the proportion of time
25 spent undertaking each of the behaviours. However, analysis of the data is problematic (Ramson and
26 Cade 2009). Even if the same animal is repeatedly sampled (for example on different days), the
27 averages and some measure of variability or precision are normally calculated for each of the
28 behavioural categories included in the ethogram separately (Ramson and Cade 2009). While
29 inferential statistics could be used to calculate significant differences between individuals in terms of
30 the occurrence of a specific behaviour, there are problems with the independence of these data
31 both in terms of repeated measures, and because all behaviours must sum to 1 as they are mutually
32 exclusive— see Aitchison 1986 and Underwood 1996 for more details about these points). Even if
33 such strict limitations on data analysis are relaxed, then this still only indicates whether animal X
34 conducts behaviour A more or less frequently than animal Y.

35 Because of these issues, it would be preferable to use a multivariate method to analyse the
36 overall behaviour of individuals (defined as all behaviours in the activity budget e.g. Mielke and
37 Berry 2007) and to compare it to other individuals for whom identical data are held. Principal
38 Component Analysis (PCA), and associated plotting of resultant components in 2 or 3 dimensions, is
39 one possible method (i.e. biplots, where any given case is plotted against the first two principal
40 components). This can give an indication of how different animals behave, based on all the
41 behaviours examined. However, several limitations to this technique exist. It is generally

42 recommended that the case to variable ratio for PCA is > 3:1, that is the number of observed animals
43 should be >3 times the number of behaviours examined (Tabachnick and Fidell, 1989), and that
44 ideally the number of cases should be high (> 300, Comrey and Lee 1992). Given that most
45 ethograms include a large number of different behaviours, and the number of animals studied is
46 often small, these limitations are significant. It is possible to use replicate sampling of the same
47 animal to boost the number of cases (using each replicate sampling period as a separate case), but
48 differences are likely to occur in behaviours based on factors such as time since eating, proximity of
49 other individuals of the same or opposite sex, hormonal changes or seasonal changes. Furthermore,
50 with traditional PCA techniques, it is not possible to determine whether differences in behaviour are
51 statistically significant or not (despite techniques such as concentration ellipses, which do not give a
52 good indication of statistical differences). Theoretically, if a biplot indicates clustering of cases from
53 one animal, and distinct, separate clustering of cases from a second animal, then they are likely to
54 be different, but, in practice, points are often interspersed and overlap with one another for the
55 reasons mentioned previously. As such, judging differences in behaviour becomes very subjective
56 (Gabriel 1971).

57 A method of combining inferential statistics with PCA has recently been developed, based on
58 constructing bootstrapped confidence intervals (or confidence radii since precision is calculated in
59 three dimensions) for each case in the PCA (Catlin-Groves et al. 2009). Because this technique
60 calculates the precision of the mean using confidence intervals, many limitations of PCA, such as the
61 case to variable ratio are less important, since lack of precision on the PCA axes is indicated by
62 increased confidence intervals. Furthermore significant differences can be inferred on the basis of
63 whether confidence radii overlap (Catlin-Groves et al. 2009). As such, the technique should be
64 beneficial for application to activity budget behavioural data collected through ethogram studies.

65 Here we develop the framework for applying this technique to activity budget studies, and
66 show the results of its application to four studies (captive and non-captive mammals, and
67 invertebrates) that indicate its potential broad application.

68 **Field data collection**

69 *Tigers in captivity*

70 Data collection took place at West Midland Safari Park in Bewdley, Worcestershire, UK (52°22'51" N,
71 2°17'06" W). In total, four Bengal tigers (*Panthera tigris tigris*) were studied, in two pairs. Each pair
72 cohabited permanently, and was moved around a number of enclosures on a day-by-day basis. The
73 first pair (tigers 1 and 2) was an unrelated male-female pair and the second pair (3 and 4) was a male
74 – female sibling pair. The enclosures in which the tigers were studied contained trees and a dual
75 layered platform in the centre of the compound. One of the enclosures also contained a small pool.

76 Data were collected in 1 h or 2 h periods, with behaviours recorded on an ethogram (Table
77 1) at 30 second intervals. In total 12 h of data were collected for each tiger (with a data point
78 collected from pairs of tigers simultaneously), giving 1440 ethogram observations per tiger.

79

80 *Elephants in a nature reserve*

81 This study was conducted at the 73.6km² Pongola Nature Reserve in South Africa (27°28'18"S,
82 31°56'49"E). Data were collected on five adult males using instantaneous scan sampling at 5 min
83 intervals (as per Altmann 1974). At each scan the behaviour of each elephant was recorded using the
84 behavioural categories listed in Table 2. Data from each male was collected until the male's
85 behaviour could no longer be accurately visually identified using binoculars. In total 154 data
86 collection points were collected for the five elephants, with a minimum of 22 ethogram observations
87 per individual.

88

89 *Dogs in rescue shelters*

90 Dogs were studied at Cheltenham Animal Shelter, Gloucestershire, UK (51° 54' 50.84" N, 2° 4' 59.51"
91 W). Dogs had already been assigned a traffic light coding of behaviour with red dogs being
92 aggressive and green dogs being friendlier and with fewer behavioural problems. This coding was
93 decided from a preliminary behaviour assessment by the shelter staff when the dogs entered the
94 shelter. Three red dogs and three green dogs were observed while being exercised in the shelter's
95 paddock. Each dog was observed three times for a total of 20 minutes, and behaviours noted every
96 20s from the list in Table 3. In total 180 ethogram observations were collected from each dog.

97

98 *Shore crab behaviour to a simulated predator*

99 Crabs were collected from a rockpool at Crantock Beach in Cornwall, UK (50°24'20" N, 5°07'51" W).
100 The rockpool was ~ 2.5 m above chart datum. For each trial, three crabs were transferred to a 1 m
101 diameter plastic experimental arena (filled with 10 cm depth of freshly collected seawater), located
102 *in situ* next to the rock pool, and allowed to acclimatise for 1 h before being observed for 10 mins.
103 During this 10 min period, crab behaviour was recorded every 30 s from the list of behaviours in
104 Table 4. Crabs were placed in groups of either three adult crabs (carapace width > 40 mm) or
105 juvenile crabs (carapace width > 20 but < 40 mm) and for each group, they were either left free from
106 visual disturbance over the 10 min period or were presented with a shadow of a predator (a
107 silhouette of a seagull) for 10 s at 60 s intervals. In total 24 crabs were used, hence each of the four
108 treatments (adult or juvenile, in the presence or absence of a visual predator stimulus) was
109 replicated twice. Each crab had 20 ethogram observations. After the study, crabs were released back
110 into the rockpool from which they came. In no cases were crabs removed from their natural
111 environment for more than 2 h.

112 **Statistical methods**

113 The bootstrapped PCA process was derived from that described in Catlin-Groves et al. (2009) and
114 slightly modified here for use on behavioural datasets. The code runs in the R statistics environment
115 (R Core Development Team 2011) and is available as supplementary material to this paper, along
116 with a sample dataset used in this study (the tiger dataset).

117 For each analysis, a frequency distribution table was set up for all cases using a spreadsheet.
118 A unique classifying number for each behaviour in the ethogram was assigned (e.g. for Table 1, '1'
119 would be assigned to feeding, '2' to foraging and so on). This classification number was typed into
120 the spreadsheet in a vertical column, with the number of entries corresponding to the percentage
121 frequency of that behaviour. For example, if behaviour 1 occurred 32% of the time, the term '1'
122 appeared in the first 32 rows of the spreadsheet. As such, each case is inputted in separate columns,
123 and behaviours indicated in rows 1-100. The number of different columns was equal to the number
124 of cases being considered within a specific analysis. The term 'case' is defined by the user. In most
125 studies here, it is the combined ethograms from any individual animal, over all the sampling periods,
126 but could be combined data from ethograms for an individual on days it had been fed, as compared
127 to days it had not been fed, for example, or multiple individuals within a particular category such as
128 sex. This classification of 'case' is considered in greater detail in the discussion and examples of
129 different classifications of case are given in the results. The conversion of behaviour into percentages is
130 to ensure that there were always 100 data points in each sample, and allow consistent rules to be
131 formulated (such as the size of the subsample for bootstrapping) to apply the technique generally to
132 behaviours where the number of observations can vary (as per the studies considered here).

133 From each case, 100 points were randomly taken, with replacement, to obtain a sample of
134 the behaviour (the use of 100 points – with replacement – from 100 does not imply all points are
135 sampled each time, and is the overwhelming consensus of sample size for bootstrapping in the
136 literature – e.g. Efron 1979; Crawley 2007; Martínez-Muoz and Suárez 2010). Using the 'prcomp'

137 function in R, the first three principal components of each sample were calculated and stored for
 138 each case, and the process repeated 10 000 times. A mean value of the 10 000 replicates was
 139 calculated and 95% confidence limits were calculated by excluding the highest and lowest 2.5 % of
 140 the values (Crawley, 2005). By altering this parameter to the highest and lowest 5% or 10%,
 141 confidence limits can be obtained at 90% or 80% levels, respectively. Upper and lower confidence
 142 intervals for all three of the stored principal components were averaged to give a confidence radius.
 143 The mean values of the principal components for each site were plotted in 3 dimensions and the
 144 confidence radius indicated the size of the sphere, or bubble. Plots were made using the RGL library
 145 and rgl.sphere function for R (Adler and Murdoch 2008). However, because of some issues of how
 146 principal components are calculated, the following modifications were required to produce the
 147 bootstrapped means and confidence radii.

148 Initially, the full dataset was analysed using the ‘prcomp’ function to give a baseline value for
 149 each case. For each replicate run of the bootstrapped principal components (where n = 100; but
 150 sampled with replacement), the full dataset (where n = 100; but without replacement) for each case
 151 was also analysed, essentially doubling the cases in replicate run. By calculating a vector to
 152 transform each point from the full dataset back to its corresponding baseline point (equation 1), and
 153 then applying the same vector to the bootstrap points (equation 2), the variability in the
 154 bootstrapped points is restricted to variation between differences in the placement of points on the
 155 initial principal component axes, and not variation between both the placement of points and
 156 alignment of principal component axes. So:

157
$$v_{[x,y,z]} = I_{[x,y,z]} - i_{[x,y,z]} \quad [1]$$

158
$$B_{mod[x,y,z]} = B_{calc[x,y,z]} + v_{[x,y,z]} \quad [2]$$

159 where v is the vector, I is the initial full data point calculated without the addition of the bootstrap
 160 points, i is the full data point calculated along with the bootstrap points, B_{mod} is the bootstrapped
 161 point modified by the vector and B_{calc} is the bootstrap point calculated directly by PCA.

162 Applying this vector also accounted for the arbitrary sign applied to the magnitude of the principal
163 component (during replicates on identical datasets, the value of a point on a principal component
164 axis could be assigned as 1 or -1). The vector transformation eliminated this problem unless the sign
165 (+ or -) of the full dataset differed from the sign of the bootstrapped dataset for the same point. If
166 this was the case, the magnitude of the vector in this dimension was ~ 2 x that of the magnitude of
167 the value of the full dataset point. To account for this problem, if the magnitude of the vector
168 exceeded 1.2 x that of the magnitude of the value of the full dataset point, the magnitude of the
169 vector in this dimension was calculated by adding the two points (equation 3) and then subtracting
170 the calculated bootstrap value from the vector (equation 4).

$$171 \quad v_{[x,y,z]} = I_{[x,y,z]} + i_{[x,y,z]} \quad [3]$$

$$172 \quad B_{mod[x,y,z]} = v_{[x,y,z]} - B_{calc[x,y,z]} \quad [4]$$

173

174 The value of 1.2 x the magnitude as the demarcation between equations 1 and 3 being applied was
175 previously been shown to be suitable, and sensitivity analysis of the results indicate that values
176 between 1 and 1.5 do not cause changes in output (Catlin-Groves et al. 2009).

177

178 **Results and Discussion**

179 *Tigers in captivity*

180 Using the standard 'prcomp' function on the full data set, the first three principal components were
181 shown to explain 99.0% of the total variance of the data set. Analysis of the four tigers showed that
182 the two females (2 and 4) had overlapping bubbles indicating that their behaviours were not
183 significantly different from each other. The two males had bubbles which also overlapped, but tiger
184 3 had a significantly different behaviour from tiger 4, but not from tiger 2 (Figure 1). Tiger 1 showed

185 significantly different behaviour compared to both females. The male tiger 3 was more similar in
186 behaviour to the two females than the male tiger 1 – which spent a considerable less time pacing
187 than the other three individuals. Indeed, Tiger 1 was recorded pacing on average 2.75 times per day,
188 compared to an average of 65.5 times per day with the other male, Tiger 3. While little work has
189 been conducted on sex specific behaviours in captive carnivores, some studies (e.g. Renner and
190 Lussier 2002) have found sex specific differences to certain aspects of captive carnivore behaviours.
191 The results from this study provide some support for sex specific differences in captive tiger behaviour,
192 but also indicate that variability between individuals may be as important as sex based differences.

193

194 *Elephants in a nature reserve*

195 Using the standard 'prcomp' function on the full data set, the first three principal components were
196 shown to explain 92.1% of the total variance of the dataset. Analysis of the five bull elephants'
197 activity budgets using the bootstrapping methods showed no significant differences between
198 elephants at the 95 % confidence level, as there is overlap between all of the coloured bubbles that
199 represented the individual elephants (Figure 2a). Such a lack of difference in activity budgets may be
200 unsurprising, given that activity levels in the African savannah are heavily constrained by time spent
201 resting as a means of coping with heat stress (Dunbar 1992). Moreover, elephant activity at Pongola
202 is further constrained by limited food available to this population, which far exceeds the park's
203 carrying capacity. However, the distribution of the bubbles does correspond closely to the previously
204 determined dominance hierarchy of these bull elephants (H. Zitzer unpublished data), with the left
205 most elephant being the most dominant, and the dominance hierarchy decreasing from left to right
206 (Figure 2a). Given that dominance was calculated by aggressive interactions, and these data
207 presented in this study are from activity budgets (where dominance interactions are largely absent),
208 such a correlation of results is a good indication that the technique is incorporating many aspects of
209 the elephants' behaviour.

210 The plot of all five bull elephants can make determining significant differences between non-
211 adjacent individuals difficult. However, pairwise comparisons can also be plotted, without the
212 analysis being rerun. To minimise type I errors of pairwise comparisons, it is logical to examine the
213 furthest apart individuals first (here elephants 1 and 5), as per the procedure in standard post-hoc
214 tests such as Student-Newman-Keuls (SNK) tests. In this case, while no significant differences occur
215 at the 95% confidence level (Figure 2b), differences do occur at the 90% confidence level between
216 the overall activity budget of elephants 1 and 5 (Figure 2c). From an examination of the activity
217 budget data, it can be seen that the key differences in behaviour are an increase in resting and
218 feeding, and a decrease in moving in the most dominant elephant, as compared to the least
219 dominant (Elephant 1 – movement = 43%, resting = 26%, feeding = 25%; Elephant 5– movement =
220 58%, resting = 17%, feeding = 15%). The differences in activity budget between the highest and
221 lowest ranking male are in line with previous field observations of these elephants. The dominant
222 male spent nearly all of his time travelling with the larger of the two female herds. As he constantly
223 had access to females, the dominant male travelled less and spent more time resting and feeding
224 with the females. The subordinate male spent a significant amount of time alone wandering
225 between the two female herds attempting to gain access to the females and as a result spent
226 significantly more time moving than the dominant male (K. Slater and H. Zitzer unpublished data).

227

228 *Dogs in rescue shelters*

229 Using the standard 'prcomp' function on the full data set, the first three principal components were
230 shown to explain 92.1% of the total variance of the data set. The bubble plot displayed some
231 significant differences between dogs (Figure 3a). The clustered group of three bubbles represent the
232 red dogs, and the three separated bubbles represent the green dogs. There is a clear significant
233 difference between all three green dogs in relation to one another and each of the red dogs,
234 indicating that their initial behavioural classification could also be determined by activity budget

235 ethograms. By redefining the classes used here, it was also possible to determine if differences occur
236 between the red and green dogs studied in general. By combining all data on the three red and three
237 green dogs, the process can be rerun. This is case, pooling the data in this way demonstrates that
238 there is not an overall significant difference between the red and green dogs, despite each individual
239 green dog being different from all red dogs (Figure 3b), although again, a significant difference
240 occurs at 90% confidence (Figure 3c). As with the tiger data, such a response indicates that variability
241 between dogs (especially the green classified dogs) can be large. In this case, differences in green
242 dog behaviour are larger than between red dogs. This may be explained by the fact that red dogs are
243 classified by aggressive characteristics – hence all behave in an aggressive manner, whereas green
244 dogs display a more natural, and varied range of domestic dog behaviours.

245

246 *Shore crab behaviour to a simulated predator*

247 Using the standard 'prcomp' function on the full data set, the first three principal components were
248 shown to explain 99.7% of the total variance of the data set. Significant differences in behaviours
249 between the treatment groups were found at the 95% confidence level (Figure 4). Juvenile crabs
250 behaved in a similar way in the absence of a predator stimulus to adult crabs in the presence of the
251 predator stimulus. Both juveniles and adults showed a similar response to predators (a downwards
252 movement in the plot of behaviour in Figure 4). From a re-examination of the data, this tends to
253 indicate an increase in hiding behaviour from both juveniles and adults in the presence of a predator
254 (from 13 to 37 % of the time in mature crabs and from 47 to 75 % of the time in juveniles).

255 Differences in behaviour of crustaceans, especially in regard to life- and moult-cycle stage, are well
256 classified, with reduced locomotion and feeding activity at the most vulnerable stages (e.g. Lipcius
257 and Herrnkind 1992), hence while both adult and juvenile respond to a predator stimulus by hiding,
258 they start from different baseline activity behaviours.

259

260 *The statistical methods*

261 The technique of bootstrapping PCA analysis works well on the examples of activity budget /
262 ethogram-recorded behaviours studied here. The technique is flexible as regards: the number of
263 samples taken per animal, the confidence level examined, and, to a large extent, the definition of
264 'case', which could be an individual animal, or a group of animals (of the same sex, age group or any
265 other logical classification). However, there are some potential considerations and
266 recommendations for the application of the technique.

267 Firstly, the number of ethogram recordings used (or the sample size) must be large enough
268 to provide a good estimate of the activity budget of the animal studied. While the conversion of
269 different behaviours to percentages (hence the effective sample size is always 100) will not affect
270 the confidence interval size of a bootstrap method, clearly, limited recording may not capture the
271 full behaviour of the animal, as such, it is best to use similar sample sizes for different animals in the
272 study and to report the sample sizes used in the methods or results.

273 Secondly, all the data sets considered here had very large proportions of variability
274 explained by the first three principal components (> 90% in all cases). This means that the positions
275 of the bubbles on three dimensional plots are accurate simplifications of the multivariate complexity
276 inherent in the original data. If the proportion of variability explained by the first three principal
277 components decreases, the number of dimensions required of the plots needs, theoretically, to
278 increase – although this would make visual interpretation of the data very difficult. As such it is
279 recommended that this technique only be used where > 90% of the variability in the data is
280 explained by the first three principal components (this figure also follows standard practice
281 recommendations for biplots given in Crawley 2007).

282 Thirdly, the technique will naturally face some of the disadvantages of all confidence interval
283 methods as compared to inferential statistical hypothesis tests (Lanzante 2005). For example,
284 confidence interval estimation for univariate methods is not as powerful as equivalent t-tests or
285 ANOVA, at least when the data fulfil parametric assumptions. However, following the procedures
286 derived for ANOVA post-hoc tests, which involve testing the most different cases first, reduces the
287 number of pairwise comparisons which need to be made (see elephant example above).
288 Furthermore, corrections to eliminate type I error could easily be made by increasing the level of
289 significance from 95%, as per Bonferroni corrections or that occur in the standard Tukey test,
290 although this should be undertaken with caution since many authors advise against such
291 modifications due to the unproportional risk of type II error over minimising type I error (e.g.
292 Underwood 1996). Whether or not such changes to confidence limits need to be made depends on
293 the study in question, and whether interpretation of results is most sensitive to falsely detecting
294 differences, or not detecting real differences. While these modifications can help prevent issues of
295 type I error, the problems of pooling estimates of variability to a common standard deviation, which
296 can result in type 2 errors (the type most frequently found with the use of confidence interval
297 analysis - Lanzante 2005) do not apply to bootstrapping processes, where confidence intervals are
298 estimated directly from the data, and do not require an estimation of standard deviation.
299 Furthermore, the bootstrapping process does not necessarily result in symmetrical confidence
300 intervals around the mean, making the technique robust to the assumptions for parametric statistics
301 such as normally distributed data. Therefore, in many ways, the bootstrapping method detailed here
302 is more robust than many statistics for hypothesis testing, which require the homogeneous standard
303 deviations and normally distributed data between cases (Underwood 1996).

304 Finally, the issue of selecting a 'case' is not as advanced as for some statistical techniques. In
305 normal PCA, a case would correspond to a single observation period. Here, multiple observation
306 periods of a single individual can be combined as a case, as can multiple observation periods of many
307 individuals within a group (providing the replication of directly observed behaviour proposed by

308 Dawkins 2007). While this provides a flexible framework for hypothesis testing, a parallel can be
309 drawn with nested designs in general linear models. Nesting hierarchical responses (i.e.
310 observations of the same individual are nested within each individual, individuals of the same sex
311 are then nested within sex), rather than simply combining responses across all levels would,
312 potentially, allow differences in individuals, as well as differences between higher level 'cases' to be
313 determined in a single analysis, and allow an understanding of where the greatest variability lies (i.e.
314 between a behavioural category, between individuals or between replicate measures of the same
315 individual). However, such an approach would not present data in such a visually simple manner,
316 and in some cases, nesting factors within others produces less powerful inferential tests than not
317 conducting this nesting process (Hernández-Sánchez et al. 2003). A method of including nesting
318 would be a useful future improvement to this technique, however, it would also create an additional
319 level of complexity in performing the analysis, which in most cases, would not make a significant
320 difference to the outcome of the analysis.

321 The technique presented here provides an excellent framework for visualising activity
322 budget collected data and provides a novel method for determining significant differences between
323 classifications of interest within the dataset. While there are some residual issues in the application
324 of the technique, which necessitate researchers to think through analysis and interpretation of
325 resultant plots carefully, the method is a vast improvement on the statistical methods currently used
326 for such analysis.

327

328 **Acknowledgements**

329 The authors would like to thank Heinz Kohrs, director of 'Space for
330 Elephants' project and Cheltenham Animal Shelter for their assistance in with this study.

331

332

333 **References**

- 334 **Adler, D. & Murdoch, D.** 2008. *rgl: 3D Visualization Device System (OpenGL)*. R Package Version 0.77.
335 <http://rgl.neoscientists.org>.
- 336 **Aitchison, J.** 1986. The statistical analysis of compositional data (with discussion). *Journal of the*
337 *Royal Statistical Society (B)*, **44**, 1313–1325.
- 338 **Altmann, J.** 1974. Observational study of behavior: sampling methods. *Behaviour*, **49**, 227-267.
- 339 **Catlin-Groves, C.L., Kirkhope, C.L., Goodenough, A.E. & Stafford, R.** 2009. Use of confidence radii to
340 visualise significant differences in principal components analysis: application to mammal
341 assemblages at locations with different disturbance levels. *Ecological Informatics*, **4**, 147–151.
- 342 **Comrey, A.L. & Lee, H.B.** 1992. *A First Course in Factor Analysis*. 2nd Edition. Hillsdale, NJ: Lawrence
343 Erlbalm Associates.
- 344 **Crawley, M.J.** 2005. *Statistics: an Introduction using R*. Chichester, UK: Wiley.
- 345 **Crawley, M.J.** 2007. *The R Book*. Chichester, UK: Wiley.
- 346 **Dawkins, M.S.** 2007. *Observing animal behaviour: design and analysis of quantitative data*. Oxford:
347 Oxford University Press
- 348 **Dunbar, R.I.M.** 1992. Time: a hidden constraint on the behavioural ecology of baboons. *Behavioural*
349 *Ecology and Sociobiology*, **31**, 35-49.
- 350 **Efron, B.** 1979 Bootstrap methods: another look at the Jackknife. *Annals of Statistics*, **7**, 1-26.
- 351 **Gabriel, K.R.** 1971. The biplot graphic display of matrices with application to principal component
352 analysis. *Biometrika*, **58**, 453–467.

- 353 **Hernández-Sánchez, J., Haley, C.S. & Visscher, P.M.** 2003. Power of QTL detection using association
354 tests with family controls. *European Journal of Human Genetics*, **11**, 819–827.
- 355 **Lanzante, J.R.** 2005. A cautionary note on the use of error bars. *Journal of Climate*, **18**, 3699-3703.
- 356 **Lipcius, R.N. & Herrnkind, W.F.** 1992. Molt cycle alterations in behavior, feeding and diel rhythms of
357 a decapod crustacean, the spiny lobster *Panulirus argus*. *Marine Biology*, **68**, 241-252.
- 358 **MacNulty, D.R., Mech, L.D. & Smith, D.W.** 2007. A proposed ethogram of large-carnivore predatory
359 behavior, exemplified by the wolf. *Journal of Mammalogy*, **88**, 595-605.
- 360 **Martin, P.R. & Bateson, P.** 2007. *Measuring Behaviour: an Introductory Guide*. 3rd Edition.
361 Cambridge: Cambridge University Press.
- 362 **Martinez-Munoz, G. & Suarez, A.** 2010. Out-of-bag estimation of the optimal sample size in bagging.
363 *Pattern Recognition*, **43**, 143-152.
- 364 **Mielke, P.W. & Berry, K.J.** 2007. *Permutation Methods: a Distance Function Approach*. 2nd Edition.
365 New York: Springer.
- 366 **R Development Core Team.** 2011. *R: A Language and Environment for Statistical Computing*. Vienna:
367 R Foundation for Statistical Computing.
- 368 **Ramson, J.I. & Cade, B.S.** 2009. *Quantifying Equid Behavior: a Research Ethogram for Free-Roaming*
369 *Feral Horses*. Reston, Virginia: US Geological Survey.
- 370 **Renner, M.J. & Lussier, J.P.** 2002. Environmental enrichment for the captive spectacled bear
371 (*Tremarctos ornatus*). *Pharmacology Biochemistry and Behaviour*, **73**, 279-283.
- 372 **Tabachnick, B.G. & Fidell, L.S.** 1989. *Using Multivariate Statistics*. New York: Harper Collins.

- 373 **Underwood, A.J.** 1996. *Experiments in Ecology: Their Logical Design and Interpretation using*
374 *Analysis of Variance*. Cambridge: Cambridge University Press.
- 375 van den Berg, L., Schlider, M.B.H. & Knoll, B.W. 2003. Behaviour genetics of canine aggression:
376 Behavioural phenotyping of golden retrievers by means of an aggression test. *Behaviour Genetics*,
377 **33**, 469-483.
- 378 **Wheller, S.H. & Bennett, C.L.** 2001. Twenty-four hour activity budgets and patterns of behavior in
379 captive ocelots (*Leopardus pardalis*). *Applied Animal Behaviour Science*, **71**, 67-79.
- 380
- 381

382 Table 1. Ethogram of behaviours used for activity budget data collection of tigers

Behaviour	Description of behaviour (where required)
Eating	-
Drinking	-
Playing	Engaging in playing activities alone
Social interaction	Interacting with another tiger – either aggressive or affiliative; including grooming one another
Rolling	-
Scent marking	Spraying an object, rubbing back paws on ground or rubbing head against objects.
Walking	-
Running	-
Pacing	Repeated walking in the same pattern without an apparent goal.
Alert standing	-
Alert sitting	-
Alert laying down	Lying down with eyes open
Not alert laying down	Lying down with eyes closed
Stalking	Walking slowly with eyes fixed on one object
Grooming	-
Defecating/urinating	-
Jumping at fence	-
Vocalise	-
Other	Any behaviour that does not fit into any of the above descriptions.

383

384

385 Table 2: Ethogram of behaviours used for activity budget data collection of elephants

Behaviour	Description of behaviour (where required)
Feeding	-
Foraging	Actively searching for or extracting food items such as bark stripping
Moving	Excluding foraging
Resting	Including sleep
Socialising	Including both aggressive and affiliative behaviours
Vigilant	Elephant is standing alert
Drinking	-

386

387

388

389 Table 3: Ethogram of behaviours used for activity budget data collection of dogs (adapted from van
390 den Berg et al., 2003).

Behaviour	Description of behaviour (if required)
Barking	-
Pulling (on lead)	-
Tail wagging	-
Growling	-
Jumping Up	-
Sitting still	-
Spinning	Dog spins in circles or changes direction frequently whilst on or off the lead.
Standing upright	-
Tail erect	-
Territorial Marking	Including urination
Approach other dogs	-
Panting	-
Whining/Whimpering	-
Yawn	-

391

392

393

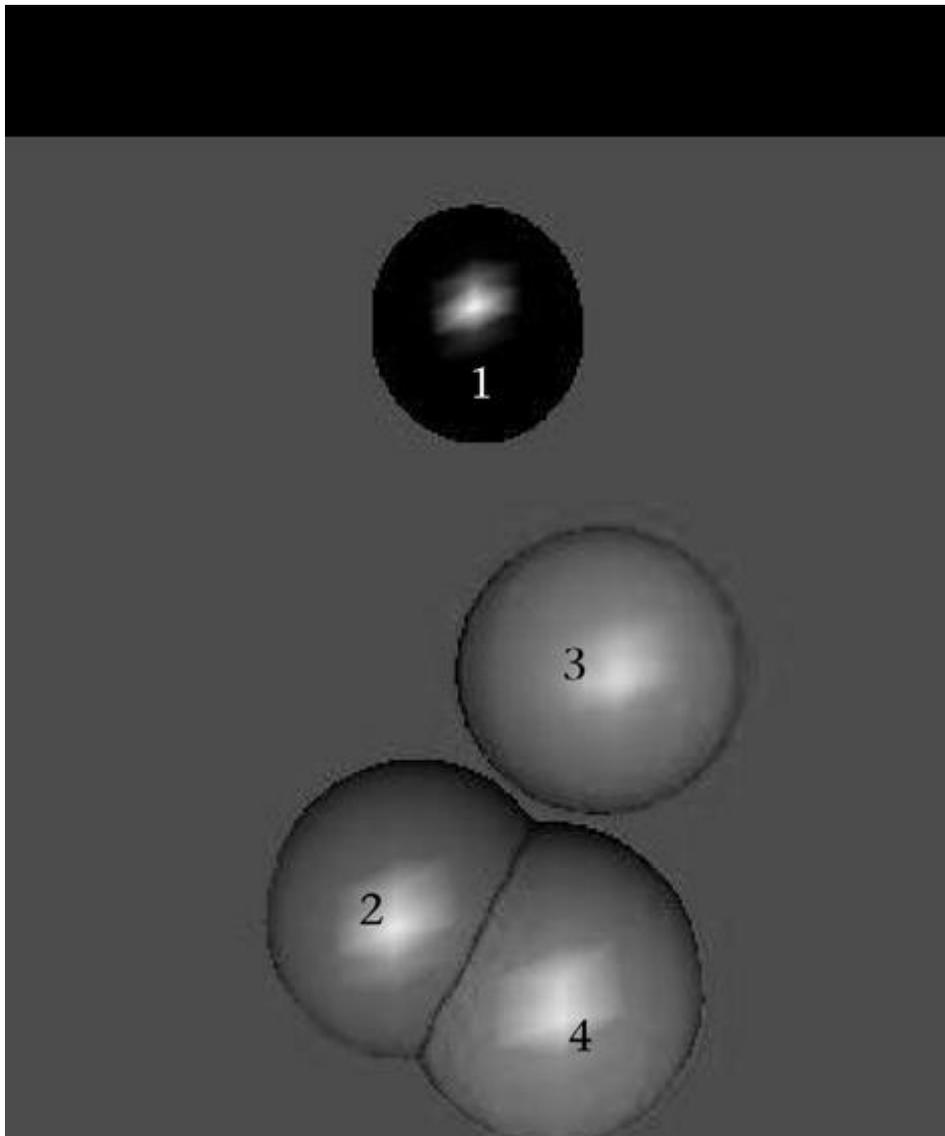
394 Table 4. Ethogram of behaviours used for activity budget data collection of shore crabs

Behaviour	Description of behaviour (if required)
Claws outstretched	-
Hide	-
Pile	Piling on top of, or forcing themselves underneath other crabs
Still	-
Quick movement	$\geq 5 \text{ cm.s}^{-2}$
Slow movement	$< 5 \text{ cm.s}^{-2}$

395

396

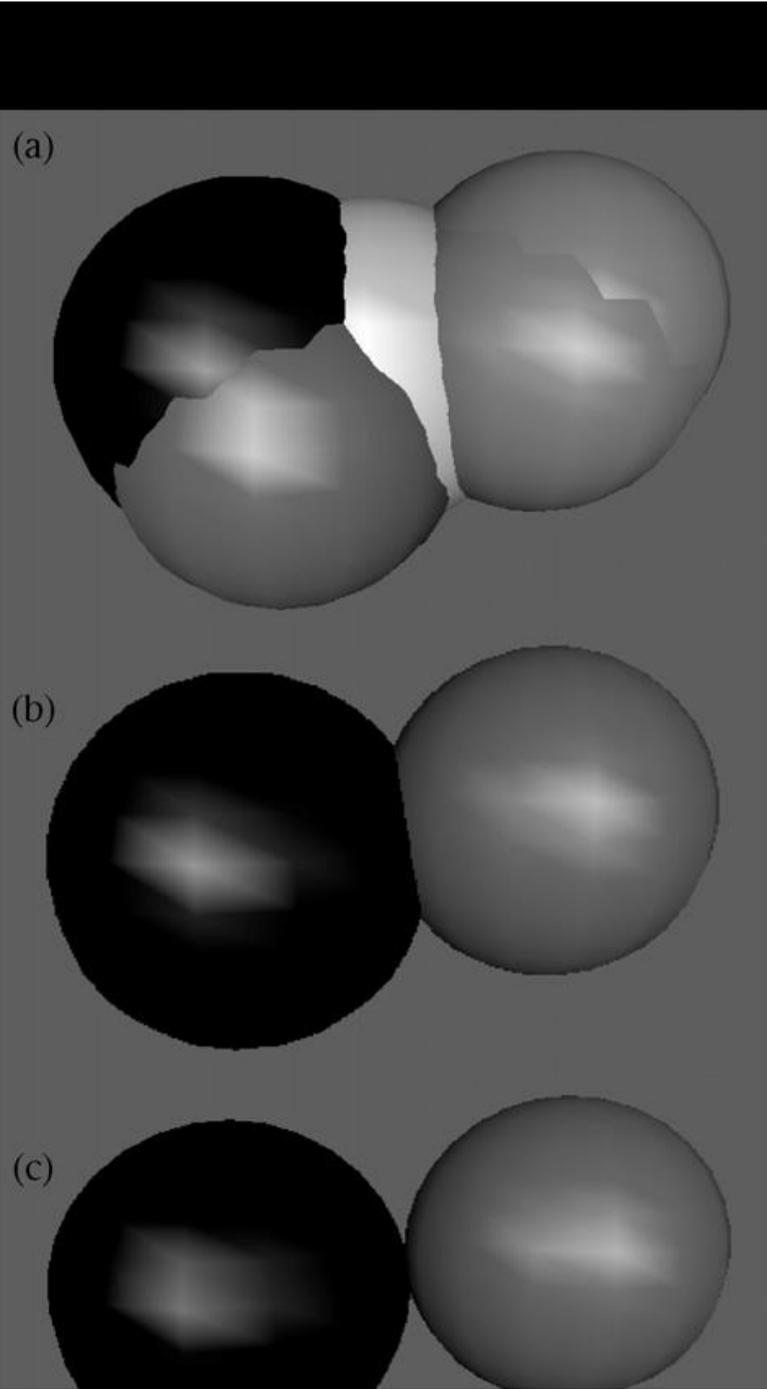
397 Figure 1. Three dimensional principal component bubble plot with confidence radii for tiger
398 behavioural data. Bubbles represent individual tigers. Tigers 1 and 3 are males and 2 and 4 are
399 females.



400

401

402 Figure 2. Three dimensional principal component bubble plot with confidence radii for elephant
403 behavioural data. (a) Each bubble represents one of the five vasectomised bull elephants, overlap of
404 bubbles indicates no significant differences at the 95% confidence level between adjacent
405 individuals. (b) Pairwise bubble plot between the most behaviourally different elephants (as
406 determined in figure 2a) at 95% confidence – overlap between bubbles indicates no significant
407 difference. (c) Pairwise bubble plot between the most different elephants at 90% confidence, here
408 no overlap of bubbles occurs, so differences can be considered significant with 90% confidence.

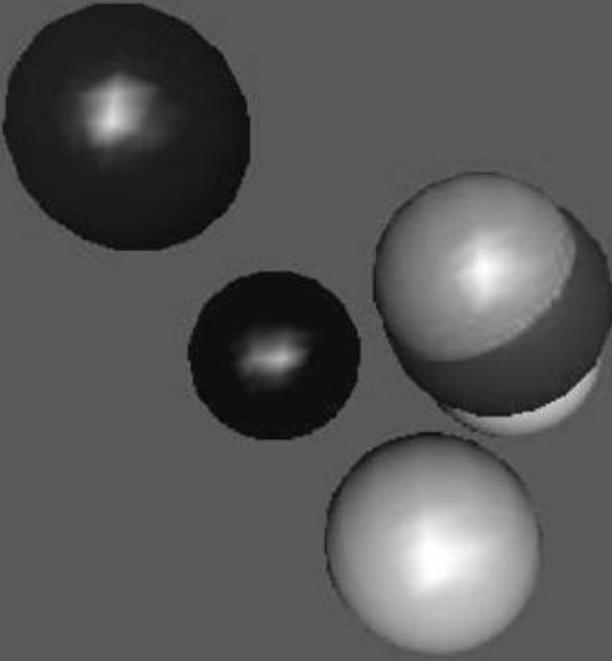


409

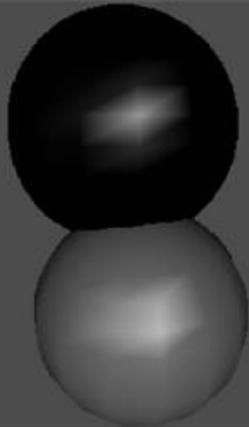
410

411 Figure 3. Three dimensional principal component bubble plot with 95% confidence radii for dog
412 behavioural data. (a) the clustered group of three dogs on the right indicate red dogs, the three
413 remaining, non-overlapping bubbles indicate the green dogs. (b) combining the data into two cases,
414 green dogs (upper bubble) and red dogs (lower bubble) shows no overall significant difference in
415 behaviour in these classifications of dogs. (c) differences do occur at the 90% confidence level
416 between green and red dogs.

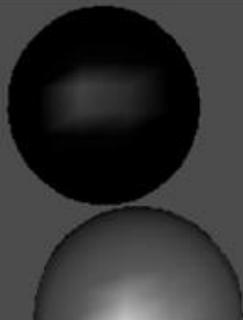
(a)



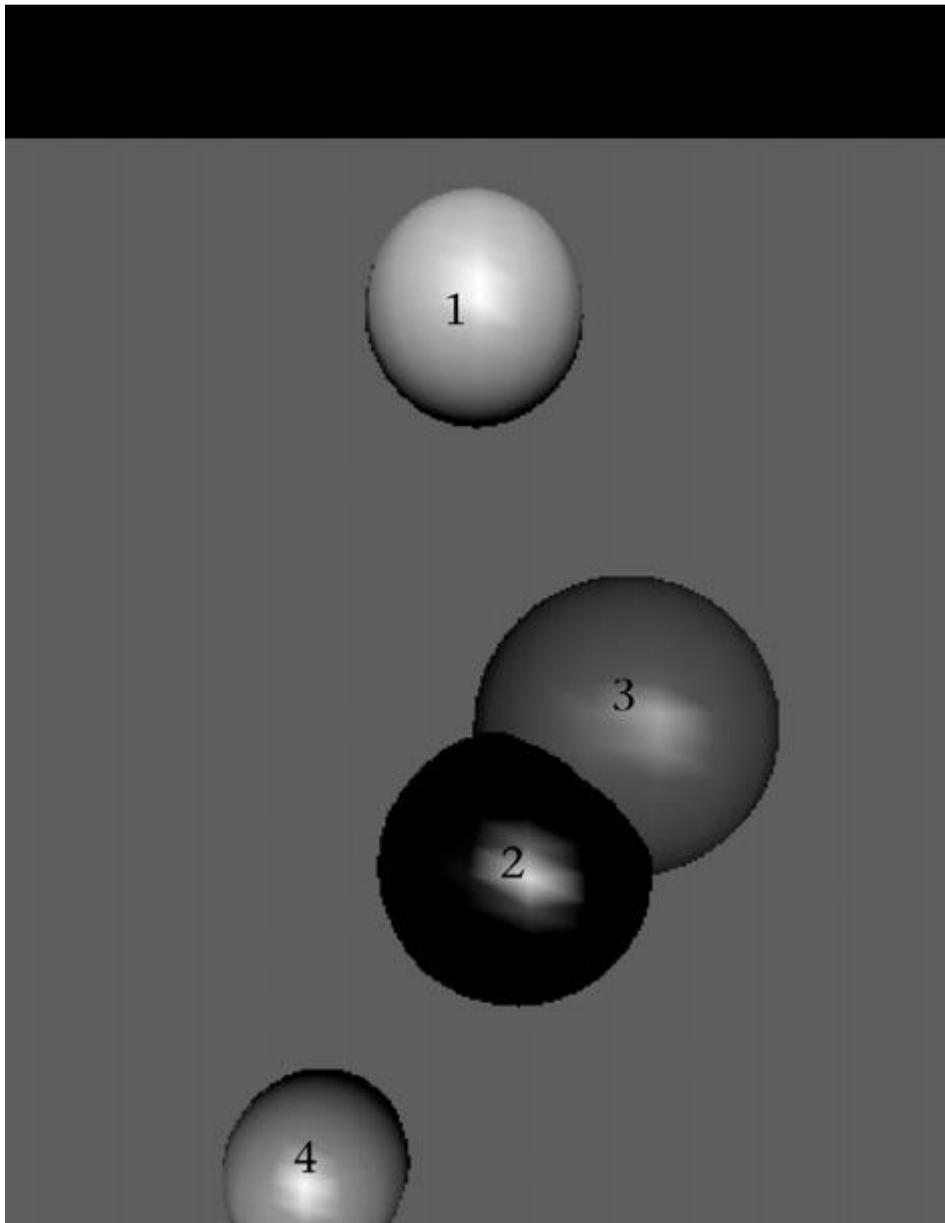
(b)



(c)



418 Figure 4. Three dimensional principal component bubble plot with 95% confidence radii for crab
419 behavioural data. Key: 1) Adult crabs in the absence of a visual predator stimulus, 2) Adult crabs in
420 the presence of visual predator stimulus, 3) Juvenile crabs in absence of visual predator stimulus, 4)
421 Juvenile crabs in the presence of visual predator stimulus.



422

423

