**A commentary on using extraction methods in data reduction analyses to assess social relationship structure in animals**

**F. Blake Mortona,c\* , Drew Altschulb,c**

a Behaviour and Evolution Research Group, Psychology, School of Natural Sciences, University of Stirling, Stirling, U.K.

b Department of Psychology, The University of Edinburgh, Edinburgh, U.K.

c Scottish Primate Research Group, U.K.

**Word Count (excluding references):**

**\*Correspondence:**

F.B. Morton, Psychology, School of Natural Sciences, University of Stirling, Stirling FK9 4LA, U.K. E-mail address: [morton.blake@gmail.com](mailto:morton.blake@gmail.com) (F. B. Morton).

**Keywords:** capuchin monkey, *Sapajus apella*, principal components analysis, factor analysis, Kaiser’s criterion, scree test, parallel analysis

**Introduction**

Comparative studies of social relationship structure (i.e. the number of dimensions and the characteristics of those dimensions) are critical to understanding animal sociality and how it evolves (refs). One problem researchers face, however, is how best to define and quantify social relationship structure such that the data are systematic and comparable across studies and taxa. Traditionally, social relationships have been described in terms of their affiliative behaviour (e.g. rates of grooming, spatial tolerance) and agonistic behaviour (e.g. rates of aggression) (Hinde 1976). More recently, Cords and Aureli (2000) proposed a 3-component model to define social relationship structure, including relationship ‘value’ (i.e. immediate benefits afforded by the relationship, such as grooming), ‘compatibility (i.e. tolerance based on partners’ shared history, such as tolerance at feeding sites), and ‘security’ (i.e. consistency and predictability in partners’ behaviour, such as rates of conflict).

Data reduction analyses like factor analysis (FA) and principal components analysis (PCA) identify inter-relationships between a set of potentially correlated variables, and cluster correlated variables into fewer discrete categories called “factors” (in FA) or “components” (in PCA) (Field 2009; Gorsuch, 1983). Because they can provide researchers with a systematic approach to categorizing different sets of behaviours (e.g. rates of grooming or aggression between group members), data reduction analyses are increasingly being used to describe social relationship structure in animals, such as Japanese macaques (*Macaca fuscata*; Majolo et al. 2010), Barbary macaques (*Macaca sylvanus*; McFarland and Majolo 2011), spider monkeys (*Ateles geoffroyi*; Rebeccini et al. 2011), capuchin monkeys (*Sapajus sp.,* formerly *Cebus apella*; see Alfaro et al. 2012; Morton et al. 2015), chimpanzees (*Pan troglodytes*; Fraser et al. 2008; Koski et al. 2012), bonobos (*Pan paniscus*; Stevens et al. 2015), and common ravens (*Corvus corax*; Fraser and Bugnyar 2010; Loretto et al. 2012). Although data reduction analyses have primarily been used to study primate social relationships, they illustrate why this approach is a potentially useful tool for studying social relationship structure across a much broader range of taxonomic groups (e.g. Fraser and Bugnyar 2010).

Before subjecting data to a data reduction analysis, one must first determine how many factors or components to extract from the analysis (Field 2009). This decision is critical given that it will influence how variables cluster together, thereby affecting the final solution (and hence) researchers’ interpretation of those results (Zwick and Velicer 1986; Ledesma and Valero-Mora 2007). Under-extraction can result in the loss of relevant information and distort the overall solution (Zwick and Velicer 1986). Over-extraction can result in some factors or components being unstable, making the overall solution difficult to interpret and/or replicate (Zwick and Velicer 1986).

Deciding when to stop extracting factors or components depends on when very little “random” variability remains in the final solution. Various cut-offs have been developed to help researchers make this decision, and involve calculating the amount of variation that is explained by each factor/component (called “eigenvalues”; Field 2009). Two commonly used methods are Kaiser’s criterion and Cattell’s scree test. Kaiser’s criterion retains components with eigenvalues >1.0; meaning, each component accounts for more variance than what is accounted for by one of the original variables (Kaiser 1960). Scree tests are a graphical technique that plots eigenvalues in a simple line plot. The number of components or factors to extract is visually estimated from the scree plot by finding the point where the line begins to level off; all components to the right of this point are considered random “noise” and should therefore be excluded (Cattell 1966).

Although each of these techniques is relatively simple to perform (perhaps contributing to their common usage by researchers), they can lead to spurious solutions. Scree plots are subjective and can lead to under- or over-extraction, particularly as the line of the plot begins to asymptote (Zwick and Velicer 1986), so it is recommended that it only be used alongside other methods.

While scree tests have maintained some utility over time, Kaiser criterion has not, a fact which has been known for some time (Revelle & Rocklin 1979). In simulation, Kaiser’s criterion’s strong bias toward over-extraction lead to a success rate of 8.77% (Ruscio & Roche 2012). The criterion failed to extract the correct number of factors in more than 90% of cases, yet Kaiser’s criterion has remained the default for popular statistics software, such as IBM SPSS.

In light of these deficiencies, many alternative methods have been developed; perhaps the most notable of which is Horn’s (1965) parallel analysis, which has been widely used, tested, and recommended by statisticians. The procedure is based on generating random eigenvalues that “parallel” the observed data in terms of sample size and the number of variables (Zwick and Velicer 1986). A factor/component is retained if its eigenvalue is greater than the 95th percentile of the distribution of eigenvalues generated from the random data (Horn 1965).

To date, most studies using data reduction analyses to describe the social relationship structure of animals have relied solely on Kaiser’s criterion (refs) or supplemented it with a scree test (refs). By contrast, parallel analysis is rarely used (Morton et al. 2015; Stevens et al. 2015). Using data from a previous study of brown capuchin monkeys (*Sapajus sp.*) as an example (Morton et al. 2015), we illustrate how one’s choice of extraction method can differentially affect the results, and thus interpretation, of social relationship structure in animals.

**Methods and materials**

*Study site and subjects*

Eighteen brown capuchin monkeys were studied at the Living Links to Human Evolution Research Centre (LL), located within the Royal Zoological Society of Scotland (RZSS), U.K. (Macdonald and Whiten, 2011). Subjects were from two breeding groups. The ‘East’ group contained four adult males, three adult females, one juvenile male and five infants (following age–sex categories in Fragaszy et al. 2004). The ‘West’ group contained four adult males, three adult females, two juvenile males, one juvenile female and five infants. Infants dependent on their mothers (i.e. those less than a year old) were not included as study subjects. Subjects’ ages ranged from 2 to 40 years for males (mean ± SD = 10.79 ± 8.55 years, *N* = 11) and 3 to 14 years for females (mean ± SD = 8.86 ± 3.63 years, *N* = 7). All group members were captive born except an adult male from East group, who was hand-reared, and the original wild-caught alpha male of West group; both individuals came to LL as established members of their groups. Both breeding groups were housed separately in identically designed 189 m3 indoor enclosures with natural light and near-permanent access to a 900 m2 outdoor enclosure containing trees and other vegetation, providing ample opportunity to engage in natural behaviours. All monkeys received commercial TrioMunch pellets supplemented with fresh fruits and vegetables three times daily and were given cooked chicken and hardboiled eggs once a week. Water was available *ad libitum* at all times. Further details of housing and husbandry are provided in Leonardi et al. (2010).

*Ethical note*

This study was entirely observational except for one aspect of data collection involving puzzle feeders, which were placed within the monkeys’ outdoor enclosures (see ‘Behavioural sampling’). Subjects could interact freely with the puzzle feeders, which were made entirely of non-hazardous material. The feeders provided a source of food snacks (raisins) and enrichment to subjects. This study was approved by Edinburgh Zoo and the ethics committee of the Psychology Department at the University of Stirling, and complied with the ASAB (2012) Guidelines.

*Behavioural sampling*

Behavioural data come from a previous study by Morton et al. (2015). Fifty-four hours of focal observations were recorded between May and August of 2011, totalling 3 h per individual. Behaviours (Table 1) were recorded daily per focal monkey for 10 min. Monkeys were sampled evenly between 0900 and 1730 hours. Incidences of aggression, coalitions, scrounging and food sharing were recorded continuously; all other behaviours were recorded at 1 min intervals using point sampling (Martin and Bateson 2007). In each point sample, group members within two body lengths from the focal were recorded. The total number of sampling points was the same for all subjects.

Between 15 May 2011 and 8 June 2011 five puzzle feeders were introduced to the outdoor enclosures of each group. Monkeys could freely interact with the feeders. Each feeder was made out of a cylindrical piece of white piping (length: 76.2 cm; diameter: 5.08 cm), with approximately 8–10 holes drilled into it (see Appendix Fig. A1 in Morton et al., 2015). Feeders were attached vertically to trees, 2–10 m apart. For each feeder, the bottom of the pipe was left open while the top of the pipe was closed. Ten paper packets, each containing five raisins, were placed in the top portion of each feeder, and wooden sticks were inserted into the holes of the pipes to prevent the packets from falling out from the bottom. The packets dropped freely from the pipe once all the wooden sticks had been removed by the monkeys. Feeders were introduced 4 days a week for approximately 30 min each day or until all of the puzzle feeders had been solved. During sessions, all instances in which a monkey approached another monkey at a feeding site were recorded, as well as the behavioural response of the receiving monkey (i.e. by avoiding or staying). East group underwent 8 sessions and West group underwent 10 sessions.

Following previous studies of study social relationship structure (e.g. Rebeccini et al. 2011; Koski et al. 2012), a set of behavioural measures (Table 1) were calculated as events per monkey dyad and subjected to a principal components analysis (PCA) with varimax rotation. Overall mean numbers of social dyadic interactions are provided in Appendix Table A1 in Morton et al. (2015).

*Statistical analyses*

All analyses were conducted in the R programming language, using the psych package (Revelle 2015). Following previous studies, component loadings greater than |0.4| were considered salient (e.g. Koski et al. 2012). Components with high loadings (i.e. |0.7|) and/or those with four or more loadings greater than |0.4| were considered robust (Guadagnoli and Velicer 1988). Seventy-three dyads and 10 behavioural measures were entered into each PCA, which meets previous recommendations for having a fixed ratio of at least 5 between the sample size and number of variables (Gorsuch 1983). Components with high loadings (i.e. |0.7|) and/or those with four or more loadings greater than |0.4| were considered robust (Guadagnoli and Velicer 1988). The 3 component PCA solution comes from Morton et al. (2015).

**Results**

*Determining the number of factors or components*

A variety of methods can and should be used to determine the number of factors/components since few datasets yield an immediate and clear solution. The “nfactors” function of the psych package produces comprehensive preliminary statistics and an accompanying chart (Figure X). Some of the included statistics can and will be inconclusive (e.g. Root Mean Residual), but for this data set, 2 factors/components were the most supported (Very Simple Structure, Complexity, Empirical BIC).

A separate function - “fa.parallel” – conducts a parallel analysis, and plots the results in the style of a Scree plot, along with Kaiser’s Criterion (Figure XX). As initially described, parallel analysis was not appropriate for FA, only PCA (Horn 1965). While FA and PCA often produce very similar solutions in practice, the underlying matrix algebra differs such that when each procedure is repeated, as in parallel analysis simulations, the results can also differ considerably. Thus, parallel analysis must be adjusted to support FA (Revelle 2015). Since we chose to use PCA for data reduction, this is not an issue, but we wish to highlight the nuance of all such analyses.

Kaiser’s criterion suggested that 3 components ought to be extracted. In contrast, both the scree plot and results of the parallel analysis (Figure XX) suggested only 2 components should be retained in the solution (Morton et al. 2015). We extracted both 2 and 3 component solutions, and compared their fit and interpretability.

*Extracted Solutions*

The 3 component solution had eigenvalues over 1.0, and explained 67.32% of the total variance. The first component was moderately correlated with the second component (r=0.493), and weakly correlated with the third component (r=-0.106). The second component was weakly correlated with the third component (r=0.01). Component 1 (Table 2) had an eigenvalue of 2.56, explained 25.6% of the variance, and was characterized by moderate to high loadings (>0.4) by behaviours reflecting the importance of the relationship in terms of social affiliation (i.e. proximity, social foraging) and direct benefits gained from this affiliation (i.e. grooming, grooming symmetry, coalitions). Component 2 (Table 2) had an eigenvalue of 2.45, explained 24.48% of the variance, and was characterized by moderate to high loadings (>0.4) from behaviours related to tolerance to approaches (avoid-stay symmetry), tolerance at feeding sites (social foraging, food sharing, food sharing symmetry), and coalitionary support. Component 3 (Table 2) had an eigenvector of 1.72, explained 17.24% of the variance, and was characterized by high loadings (>0.89) from behaviours indicating a lack of stability or predictability in the relationship (i.e. conflict and conflict symmetry).

The 2 component solution collectively explained 55.0% of the total variance. Component 1 (Table 2) had an eigenvalue of 3.72, explained 37.22% of the variance, and was characterized by moderate to high loadings (>|0.45|) on behaviours related to social affiliation (e.g. proximity, social foraging, food sharing, and grooming). Component 2 (Table 2) had an eigenvalue of 1.78, explained 17.8% of the variance, and was characterized by high loadings (>|0.869|) from agonistic behaviours, i.e. conflict and conflict symmetry. Compared to the 3-component solution, correlations between these two components were only weakly correlated (r=-0.072).

**Discussion**

The PCA solution derived using Kaiser’s criterion closely resembles the 3-component model proposed by Cords and Aureli (2000). Component 1 resembles relationship “value” (e.g. grooming and food sharing), component 2 resembles relationship “compatibility” (e.g. tolerance at feeding sites), and component 3 resembles relationship “security” (e.g. rates of aggression). By contrast, the PCA solution derived using a scree test and parallel analysis recommended two components, i.e. basic affiliative and agonistic components, which, as previously discussed, reflects a more traditional approach to describing social relationship structure (e.g. Hinde 1976).

Similar findings have recently been reported in bonobos by Stevens et al. (2015), who identified three components resembling the 3-component model proposed by Cords and Aureli (2000) when using Kaiser’s criterion and a scree test, but ended up retaining only two components (labelled ‘value’ and ‘compatibility’, respectively) based on a parallel analysis. Importantly, their findings also illustrate how a scree test, which in the current study recommended the same number of components as a parallel analysis, is not always a conservative method to use in data reduction analyses.

As previously noted, most studies have used Kaiser’s criterion, including the current study, and typically find 3 components to their subjects’ social relationships (but see Rebeccini et al. 2011). However, if one examines the individual loadings of the behaviours that are entered into each of these analyses (most of which are the same behaviours), there are striking differences in terms of how certain behaviours load onto each component. For example, in capuchins and Japanese macaques, aggression loads positively onto a component resembling relationship “security” (this study; Majolo et al. 2010), whereas in chimpanzees, Barbary macaques, and corvids, the same behaviour loads positively onto a component resembling relationship “compatibility” (Fraser et al. 2008; Fraser and Bugnyar 2010; Koski et al. 2012). Additionally, in capuchins and Japanese macaques, grooming symmetry loads positively onto a component resembling relationship “value” (this study; Majolo et al. 2010), whereas in chimpanzees and Barbary macaques, the same behaviour loads positively onto a component resembling relationship “security” (Fraser et al. 2008; McFarland and Majolo 2011). Two possible explanations may underlie these differences. First, certain behaviours like aggression and grooming symmetry may have different meanings for capuchin and Barbary macaque social relationships, which would explain why these behaviours cluster onto different components compared to solutions derived in other species. Alternatively, differences between studies may reflect instability in solutions derived using Kaiser’s criterion. This is not to say that using Kaiser’s criterion is “wrong” *per se*; rather, as noted previously, one major disadvantage to this approach is that it often leads to unstable solutions. Therefore, within the context of social relationship studies, structural differences between studies may be biologically meaningful (i.e. they could reflect differences in sociality), but they could also reflect structural instability as a result of using Kaiser’s criterion.

**Conclusion**

Collectively, the current example should serve as a cautionary note to researchers wishing to use data reduction analyses to study social relationship structure in animals. In particular, careful decisions must be made when determining how many components or factors to retain in one’s analysis. As discussed, many statisticians recommend using parallel analysis, but this method is very rarely used in studies of animal social relationship structure. Because it is a more conservative and less subjective method than other extraction techniques, parallel analysis is highly recommended for any future study using data reduction analyses to investigate social relationship structure. Doing so will greatly improve the utility (and reliability) of data reduction techniques for comparative interpretations of animal social relationships.

**References**

Alfaro JWL, Silva JDSE, Rylands AB (2012) How different are robust and gracile

capuchin monkeys? An argument for the use of Sapajus and Cebus. American

Journal of Primatology, doi: 10/1002/ajp.22007.

ASAB (2012) Guidelines for the treatment of animals in behavioural research and

teaching. Animal Behaviour, 83:301-309.

Cattell RB (1966) The scree test for the number of factors. Multivariate Behavioural

Research, 1:245-276.

Cords M, Aureli F (2000) Reconciliation and relationship qualities. In: Natural conflict

resolution (Aureli F. & de Waal F.B.M., eds.). University of California Press,

Berkley, CA, p. 177-198.

Field A (2009) Discovering statistics using SPSS, third ed. London, SAGE.

Fragaszy DM, Visalberghi E, Fedigan LM (2004) The complete capuchin. Cambridge:

Cambridge University Press.

Fraser ON, Bugnyar T (2010) The quality of social relationships in ravens. Animal

Behaviour, 79:927-933.

Fraser ON, Schino G, Aureli F (2008) Components of relationship quality in

chimpanzees. Ethology, 114:834-843.

Gorsuch RL (1983) Factor analysis, 2nd Edition. Hillsdale, NJ:LEA.

Gouadagnoli E, Velicer WF (1988) Relation of sample size to the stability of component

patterns. Psychological Bulletin, 103:265-275.

Hinde RA (1976) Interactions, relationshipos, and social structure. Man, 11:1-17.

Horn JL (1965) A rationale and test for the number of factors in factor analysis. Psychometrika,

30:179-185.

Kaiser HF (1960) The application of electronic computers to factor analysis. Ed. Psychol. Meas.,

20:141-151.

Koski SE, de Vries H, van de Kraats A, Sterck EHM (2012) Stability and change of social

relationship quality in captive chimpanzees. International Journal of Primatology, 33:905-

921.

Ledesma RD, Valeromora P (2007) Determining the number of factors to retain in EFA: an easy-

to-use computer program for carrying out parallel analysis. Practical Assessment,

Research & Evaluation, 12:1-11.

Leonardi, R., Buchanan-Smith, H. M., Dufour, V., MacDonald, C., & Whiten, A. (2010). Living together: Behavior and welfare in single and mixed species groups of capuchin (*Cebus apella*) and squirrel monkeys (*Saimiri sciureus*). *American Journal of Primatology*, *72*, 33–47.

Loretto, M.-C., Fraser, O. N., & Bugnyar, T. (2012). Ontogeny of social relations and coalition formation in common ravens (*Corvus corax*). *International Journal of Comparative Psychology*, *25*, 180–194.

Macdonald, C., & Whiten, A. (2011). The ‘Living Links to Human Evolution’ research centre in Edinburgh zoo: A new endeavour in collaboration. *International Zoo Yearbook*, *45*, 7–17.

Majolo, B., Ventura, R., & Schino, G. (2010). Asymmetry and dimensions of relationship quality in the japanese macaque (*Macaca fuscata yakui*). *International Journal of Primatology*, *31*, 736–750.

Martin, P., Bateson, P. (2007). Measuring behaviour: an introductory guide. Cambridge,

U.K.: Cambridge University Press.

McFarland, R., & Majolo, B. (2011). Exploring the components, asymmetry and distribution of relationship quality in wild barbary macaques (*Macaca sylvanus*). *PLoS One*, *6*, e28826.

Morton FB, Weiss A, Buchanan-Smith HM, Lee PC (2015) Capuchin monkeys with similar personalities have higher-quality relationships independent of age, sex, kinship, and rank. *Animal Behaviour*. Doi: 10.1016/j.anbehav.2015.04.013

Rebecchini, L., Schaffner, C. M., & Aureli, F. (2011). Risk is a component of social relationships in spider monkeys. *Ethology*, *117*, 691–699.

Revelle, W. (2015) psych: Procedures for Personality and Psychological Research, Northwestern University, Evanston, Illinois, USA, http://CRAN.R-project.org/package=psych Version = 1.5.4.

Stevens JMG, de Groot E, Staes N (2015) Relationship quality in captive bonobo

groups. Behaviour, 152:259-283.

Zwick WR, Velicer WF (1986) Comparison of five rules for determining the number of

components to retain. Psychological Bulletin, 99: 432-442.

Table 1. Behavioural measures calculated for each monkey dyad (reproduced with permission from Morton et al. 2015). \*

|  |  |
| --- | --- |
| **Behavioural measure** | **Calculation** |
| Avoid/Stay Symmetry | (no. times A approaches B) / [(no. times A approaches B) + (no. times B approaches A)] |
| Coalitions | [(no. times A supports B) + (no. times B supports A)] |
| Conflict | [(no. times A aggresses B) + (no. times B aggresses A)] |
| Conflict Symmetry | (no. times A aggresses B) / [(no. times A aggresses B) + (no. times B aggresses A)] |
| Food Sharing | [(no. times A gives to B) + (no. times B gives to A)] |
| Food Sharing Symmetry | (no. times A gives to B / [(no. times A gives to B) + (no. times B gives to A)] |
| Grooming | [(no. minutes A grooms B) + (no. minutes B grooms A)] |
| Grooming Symmetry | (no. minutes A grooms B) / [(no. minutes A grooms B) + (no. minutes B grooms A)] |
| Social Foraging | [(% of time A within proximity of B) + (% of time B within proximity of A)] |
| Spatial Proximity | [(% of time A within proximity of B) + (% of time B within proximity of A)]\*\* |

\*Modified from Rebecchini et al. (2011) and Koski et al. (2012). \*\*These calculations do not include time spent grooming or time spent social foraging.

Table 2. Varimax-rotated PCA structure of 10 behavioural measures based on Kaiser’s criterion, a scree test, and parallel analysisa

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Behavioural measure | Kaiser’s criterion | | |  | Parallel analysisb,c | |
| PC1 | PC2 | PC3 |  | PC1 | PC2 |
| Spatial Proximity | **.863** | .274 | .028 |  | **.803** | -.112 |
| Grooming | **.806** | .285 | .050 |  | **.772** | -.077 |
| Grooming Symmetry | **-.743** | .099 | .168 |  | **-.450** | .343 |
| Avoid/Stay Symmetry | .048 | **-.763** | -.036 |  | **-.507** | -.188 |
| Food Sharing | .247 | **.680** | -.202 |  | **.651** | -.120 |
| Food Sharing Symmetry | .088 | **.658** | .175 |  | **.532** | .272 |
| Coalitions | **.434** | **.658** | -.064 |  | **.771** | -.033 |
| Social Foraging | **.590** | **.607** | -.011 |  | **.846** | -.026 |
| Conflict Symmetry | .006 | .033 | **.899** |  | .049 | **.865** |
| Conflict | -.086 | -.019 | **.898** |  | -.053 | **.875** |

aSalient loadings (>|0.4|) per behaviour are in bold; PC=principal component; bA scree test, complexity, empirical BIC, very simple structure, and parallel analysis recommended that two components be extracted; c reproduced with permission from Morton et al. (2015).

Figure 1. Plotted result of the R psych package “nfactors” function. Very Simple Structure assesses

Figure 2. Results of parallel analysis, on a scree plot. Triangles indicate eigenvalues for components; X’s indicate adjusted eigenvalues for factors. Dashed lines represent random simulated eigenvalues for the corresponding factor or component procedures. The horizontal black line at 1 is Kaiser’s criterion.

Table 3. Results of Horn's (1965) parallel analysis for component extraction.\*

|  |  |
| --- | --- |
| **Component** | **Adjusted Eigenvalues** |
| 1 | **2.920683** |
| 2 | **1.202033** |
| 3 | 0.858408 |
| 4 | 0.708558 |
| 5 | 0.549795 |
| 6 | 0.568731 |
| 7 | 0.533918 |
| 8 | 0.527448 |
| 9 | 0.549315 |
| 10 | 0.568179 |

\*1,000 iterations, using the 95th percentile estimate; adjusted eigenvalues >1.0 indicate dimensions to retain; 2 components retained (boldface). Unadjusted eigenvalues are also those used for the scree test in Figure 1. Reproduced with permission from Morton et al. (2015).