Comparing the linear and non-linear Principal Component Analysis over Likert-type items: an empirical study based on balanced bootstrap

Calapez, Teresa (1st author)
Instituto Universitário de Lisboa (ISCTE-IUL), Unidade de Investigação em Desenvolvimento Empresarial (Unide-IUL)
Av Forças Armadas, s/n
1649-026 Lisboa, Portugal
E-mail: teresa.calapez@iscte.pt

Botelho, Maria do Carmo (2nd author)
Instituto Universitário de Lisboa (ISCTE-IUL), Unidade de Investigação em Desenvolvimento Empresarial (Unide-IUL)
Av Forças Armadas, s/n
1649-026 Lisboa, Portugal
E-mail: maria.botelho@iscte.pt

Ramos, Madalena (3rd author)
Instituto Universitário de Lisboa (ISCTE-IUL), Centro de Investigação e Estudos de Sociologia (CIES-IUL)
Av Forças Armadas, s/n
1649-026 Lisboa, Portugal
E-mail: madalena.ramos@iscte.pt

As part of a wider project about students' perceptions of mathematics, this study analyses two dimensions: the usefulness of mathematics and the importance of understanding the concepts. Each dimension is theoretically structured by six items, three with a positive connotation and three with a negative connotation (adaptations of Fennema and Sherman, 1976 and Kloosterman and Stage, 1992). Each item is a statement that seeks to register the intensity of agreement on a 5-point scale. Four questionnaires were constructed: in the first two (A and B) the statements are evaluated by marking the response on a line segment (Visual Analogue Scale), labelled in the extreme and at the midpoint (A) or only at the extremes (B). In the third (C), Likert-type items with five labelled points are used (completely disagree, disagree, neither agree nor disagree, agree and totally agree), while in D only the extremes are labelled. This work concerns only questionnaires type C and D. Since the variables of this type are structurally ordinal variables (Göb *et al*, 2007), the use of nonlinear principal components (e.g. CATPCA, see Linting *et al*, 2007a) is a possibility for its interdependency analysis. However, the use of linear techniques (PCA) has been widespread as a form of preferential treatment of such data. Do both techniques lead to similar results? If not, then the linearity of the treatment of such items can be questioned. Drawing on the data collected in the study mentioned, and using the balanced bootstrap as proposed in Linting *et al* (2007b), several confidence intervals are built for the objects scores and loadings, either obtained through PCA or CATPCA.The comparison of structures will also include measures as the Tucker's congruency coefficient (see Abdi, 2007). A resample method will be used to build confidence intervals for those indicators. However, to obtain confidence intervals for objects, each object has to appear as many times as the number of replicates. This procedure, called balanced bootstrap was introduced by Davison et al (1986) and is proposed in a similar context by Linting *et al* (2007b).

Participants and Procedures

The target population of the study consists of first year, first-time students, from several degree courses (Social Sciences, Management and Technological Sciences Courses) at a public university institution, whose course plan includes subjects from the scientific area of quantitative methods. The ages of the 727 participants range between 16 and 56 years, the average age is 20.9 years (SD=6.7) and the most frequent age is 18 years old. The majority of students are female (52.1%), with a Sciences' background from high school (71%) and had mathematics until their entrance in the university (91.4%). Almost half of the students are enrolled on Social Sciences courses (45.8%), while 39.3% are enrolled on Management courses and, finally, 14.9% on Technological Sciences courses; this clearly reflects the profile of the institute's undergraduate student population. As this paper deals only with types C and D questionnaires, only 364 out of the 727 students were considered. The main characteristics described above remain almost equal for this subsample. Items were randomly assigned and in the same order whatever the questionnaire type. Questionnaires were systematically distributed within classes, so that an approximate number of each type was obtained (around 180 in each type). Groups can be considered homogeneous in what concerns gender ($\chi^2_{(3)} = 5.385$, p = 0.146), age ($\chi^2_{(15)} = 14.815$, p = 0.465) and course field ($\chi^2_{(6)} = 0.303$, p = 0.999). Inspired by Linting *et al* (2007b) work, we intend to analyse the stability of the results obtained with linear and non-linear PCA. In particular, they have shown that eigenvalues and component loadings of linear and nonlinear PCA are approximately equally stable. Also, substantively, they found that results from linear and nonlinear PCA revealed to be quite similar. They also report instability due to small marginal frequencies, which led to the decision of category merging, with increased stability while maintaining overall interpretation.

The bootstrap resampling technique was first proposed by Efron (1979). The main idea is to replace the unknown theoretical distribution function of an estimator by its empirical distribution function obtained with a large number of replicates with the same dimension as the original sample (plug-in principle). This procedure allows point and interval parameter estimation along with the corresponding standard error. In 1986, Davison *et al* introduced first order balanced bootstrap as a method to guarantee that the n initial observations appear exactly B times on the B constructed replicates. For our study, this is an important feature, as we aim to obtain bootstrap confidence intervals for individual object scores as well. Moreover, not only Efron and Tibshirani (1993) found that the average performance of the balanced bootstrap was about the same as that of the simple bootstrap estimator, but also Markus (cited by Linting *et al*, 2007b) found better results with balanced bootstrap than with the unbalanced, more usual, version.

The balanced bootstrap algorithm used follows Linting *et al* and can be described in the following manner:

1. Let $X_1,\ldots,X_n$ be a sample of dimension n;
2. Let $\mathbf{v}$ be a n dimensional vector of consecutive integers;
3. Let $\mathbf{k}$ be a new vector obtained by juxtaposing B copies of $\mathbf{v}$; thus, $\mathbf{k}$ has length nB;
4. The elements of $\mathbf{k}$ are randomly permutated, generating vector $\mathbf{k}_p$;
5. The first bootstrap sample are the first n elements of $\mathbf{k}_p$; the second bootstrap sample are the next n elements of $\mathbf{k}_p$ (from the $(n+1)^{th}$ to the $2n^{th}$), and so on, until the B bootstrap replicates are built.

In our study we considered a total number of replicates, B, of 1000, as did authors in previous studies. In each bootstrap replicate a linear and a nonlinear PCA were performed, keeping in each case a 2-dimensional solution. In nonlinear PCA all input variables were defined as ordinal. Loadings, object scores, correlations between object scores of the replicates with the total sample's object scores, and eigenvalues were registered for posterior use. Also, Tucker's congruency coefficients between each bootstrap sample's loading matrix and the total sample's loading matrix were computed and registered. Tucker's congruency coefficient is a similarity measure of matrices (see Abdi, 2007); being F and G two conformable loading matrices, it can be defined as

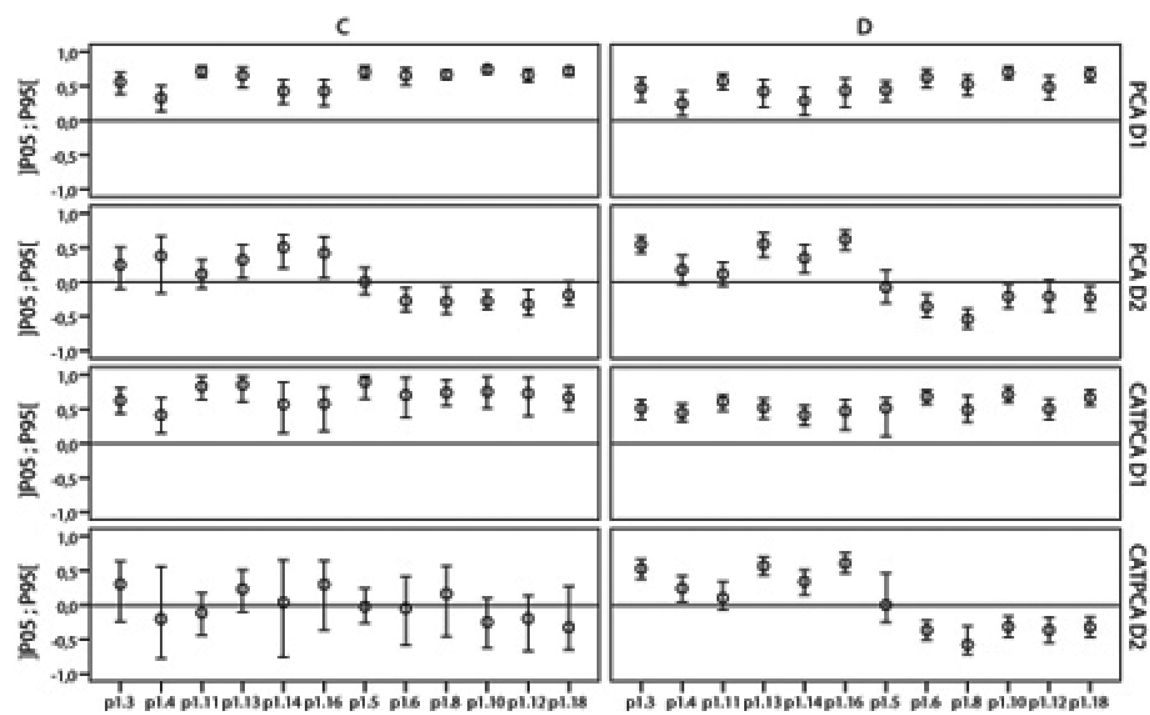
$$\phi = \frac{tr\left(F'G\right)}{\sqrt{tr(F'F)tr(G'G)}}, \qquad (1)$$

In order to guarantee comparability, dimension reflection was discretionarily performed, so that each column of the replicate's loading matrix is positively related with the corresponding column of the total sample's loading matrix.

Results

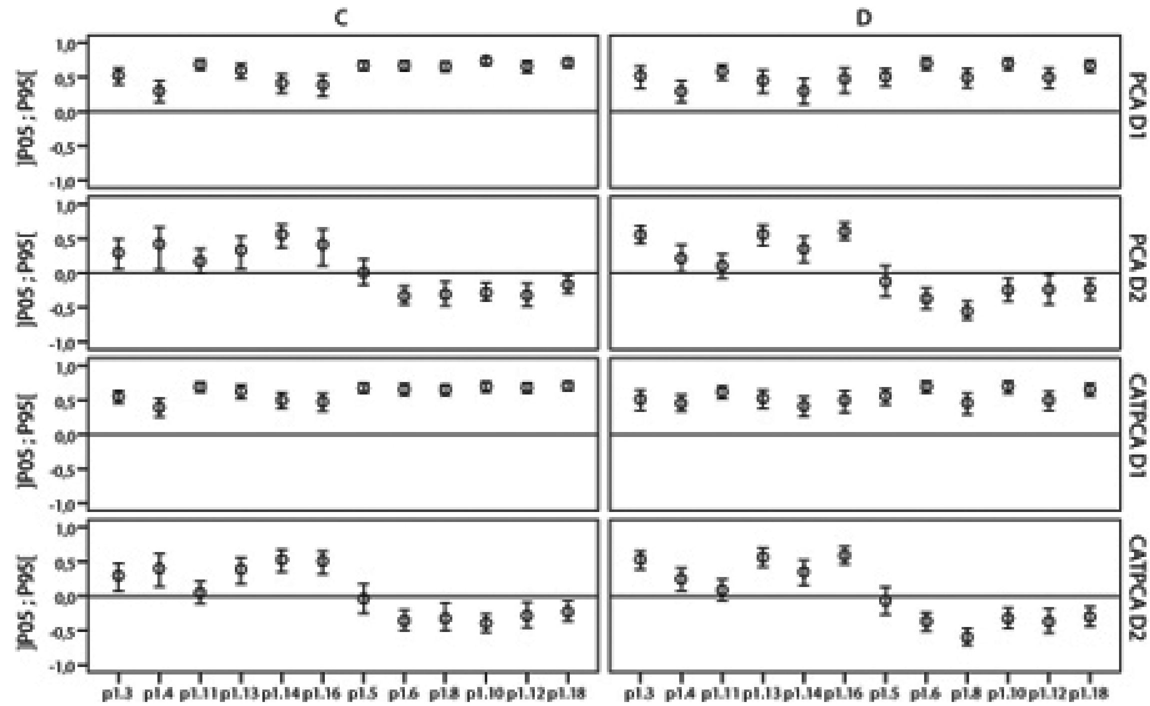
In our study, as in Linting *et al* (2007b), instability due to small marginal frequencies was observed. As in the cited study, categories were merged, revealing increased stability and correlations between object scores of each replicates and the corresponding object scores for the total sample, while maintaining overall interpretation. All bootstrap confidence intervals are of the percentile type, computed at the 90% level,]P05 ; P95[. Comparing the 90% confidence intervals for loadings of each variable, before and after category merging (Figures 1 and 2), it can be seen a decrease in CI's length for almost all variables, in both types of questionnaires and more evident in what the nonlinear PCA is concerned.

Figure 1: 95% Bootstrap confidence intervals for loadings of linear and nonlinear PCA over type C and D questionnaires (original variables)



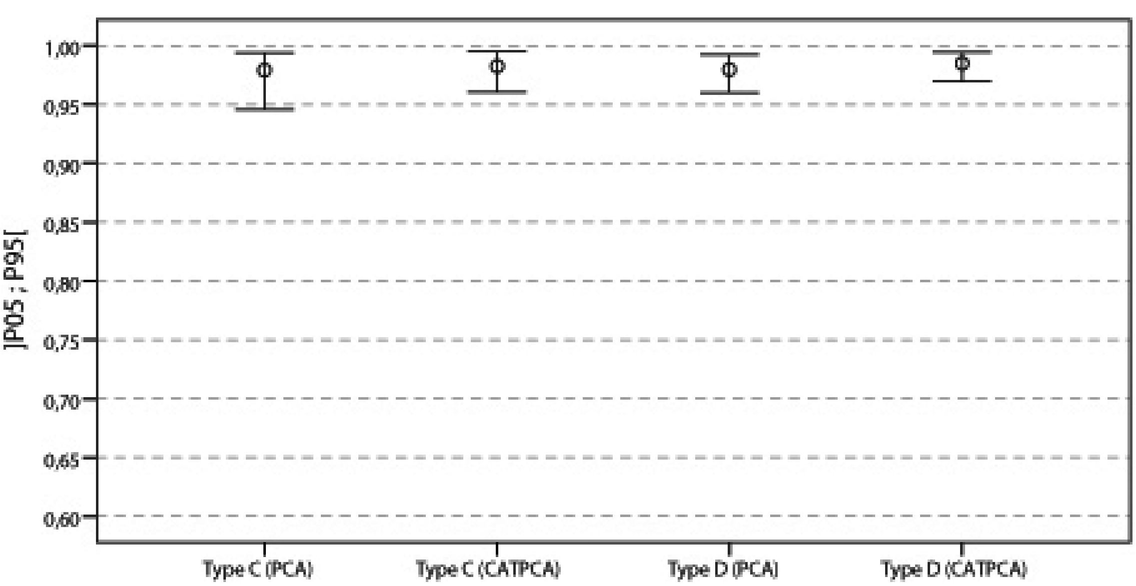
From Figure 2, it can also been seen that loading structures should be quite similar, for both questionnaires types as well as for linear and nonlinear PCA. For each replicate we have computed

Figure 2: 95% Bootstrap confidence intervals for loadings of linear and nonlinear PCA over type C and D questionnaires (recoded variables - categories merged)



Tucker's congruency coefficients between the replicate's structure and the original sample's structure, in both linear and nonlinear analysis. Confidence intervals for such coefficients are represented in Figure 3. All the 90% level CIs lie approximately between 0.95 and 1.00, thus revealing stability of the results, in both questionnaire types and in both techniques.

Figure 3: 95% Bootstrap confidence intervals for Tucker's congruency coefficients



It should be noted that, in dimension 2, confidence intervals for object scores from type D

Figure 4: 95% Bootstrap confidence intervals for first dimension's object scores of linear and nonlinear PCA, over types C and D questionnaires

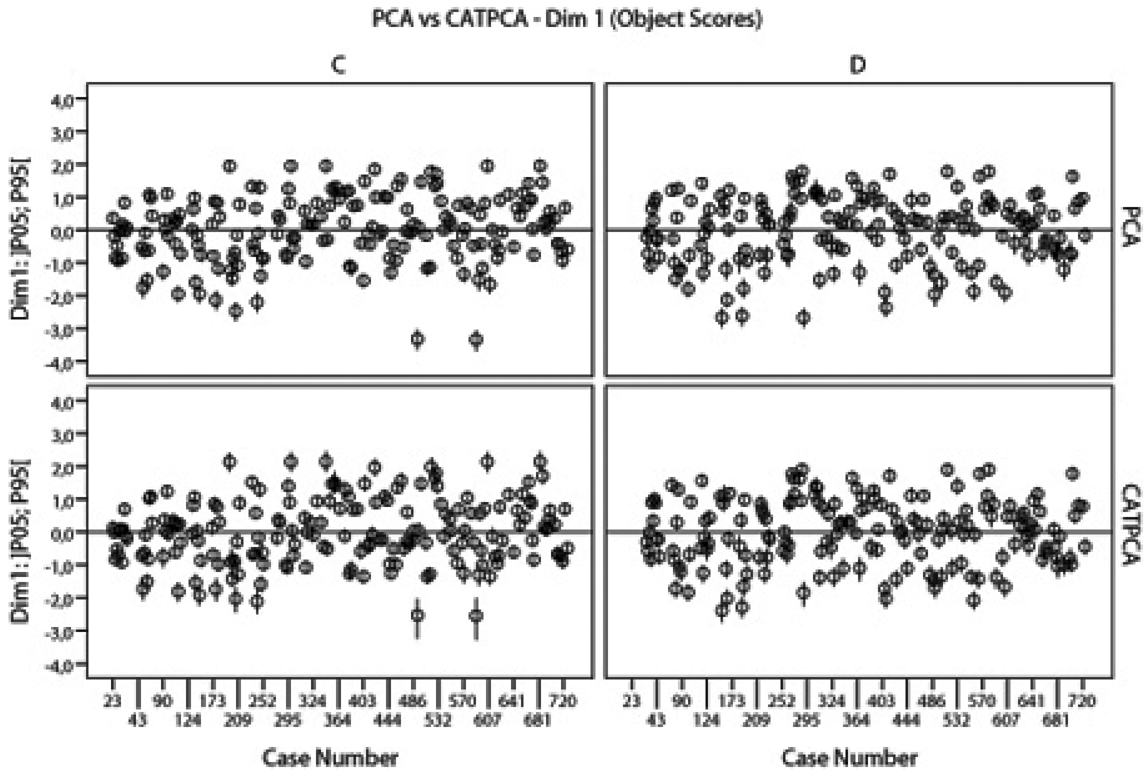
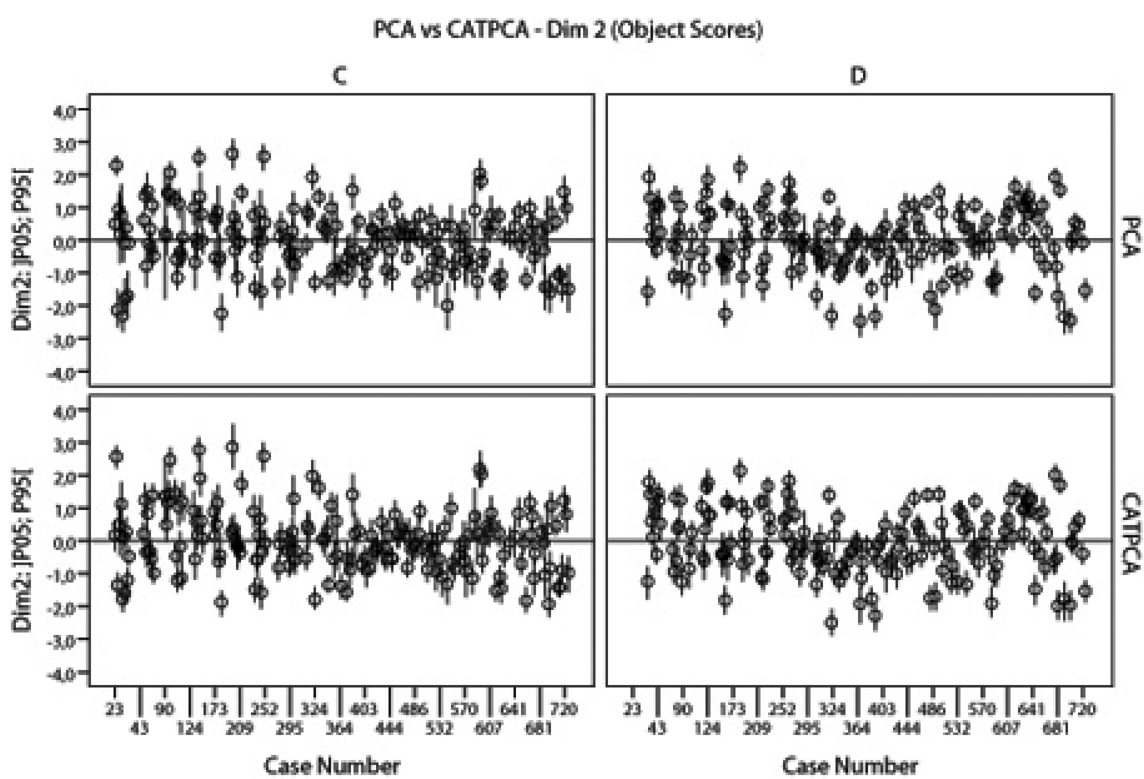


Figure 5: 95% Bootstrap confidence intervals for second dimension's object scores of linear and nonlinear PCA, over types C and D questionnaires



questionnaires seem to be tighter than those from type C questionnaires (Figure 5). Differences between questionnaires types are less evident for the first dimension (Figure 4). Confidence intervals for nonlinear PCA's object scores for the first dimension appear to be wider than those for the corresponding linear PCA's object scores. Moreover, paired samples t-tests allowed to conclude that, for the second dimension, there are no significant differences between linear and nonlinear PCAs, in what average CI's length is concerned (Table 1).

Table 1: 95% Bootstrap confidence intervals average length - paired samples t-test.

			Mean	S.d.	t	Sig
C	CI length PCA D1	CI length CATPCA D1	-0.091	0.088	-14.189	0.000
C	CI length PCA D2	CI length CATPCA D2	0.027	0.291	1.274	0.204
D	CI length PCA D1	CI length CATPCA D1	-0.031	0.075	-5.569	0.000
D	CI length PCA D2	CI length CATPCA D2	0.001	0.138	0.134	0.894

Conclusions

These results confirm the sensitivity of nonlinear PCA to the existence of residual categories and the consequent need for its merging, especially when items have all categories labeled (type C questionnaires). The good results for linear PCA in type D questionnaires - items with labels only at the extremes - suggest that the linear treatment of such items is more adequate in that case than when all categories are labeled. As a future work the authors intend to, as in Linting *et al* (2007b), proceed to the analysis of the 90% confidence ellipsis for both loadings and object scores.

REFERENCES (RÉFERENCES)

Abdi, H. (2007). RV Coefficient and Congruence Coefficient. In: Neil Salkind (Ed.), *Encyclopedia of Measurement and Statistics*. Sage.

Davison, A.C. ,D.V. Hinkley; E. Schechtman (1986). "Efficient bootstrap simulation". *Biometrika*, 73, 555-566.

Efron, B. (1979). "Bootstrap methods: another look at jackknife". *The Annals of Statistics*, 7, 1-26.

Efron, B.; Tibshirani, R. J. (1993). *An introduction to the bootstrap.* New York: Chapman & Hall.

Fennema, E. H.; Sherman, J.A. (1976). "Fennema-Sherman mathematics attitudes scales: instrument designed to measure attitudes toward mathematics". *Journal for Research in Mathematics Education*, 7(5), 324-326.

Göb, R.; McCollin, C.; Ramalhoto, F. (2007). "Ordinal Methodology in the Analysis of Likert Scales". *Quality & Quantity*, 41, pp 601-626.

Kloosterman, P.; Stage, F.K.(1992). "Measuring beliefs about mathematical problem solving". *School Science and Mathematics*, 92, 109-115.

Markus, M. T. (1994). *Bootstrap confidence regions in nonlinear multivariate analysis.* Leiden, DSWO Press.

Linting, M.; Meulman, J.; Groenen, P.; Van der Kooij, A. (2007a). "Nonlinear Principal Components Analysis: Introduction and Application". *Psychological Methods*, 12(3), pp 336-358.

Linting, M.; Meulman, J.; Groenen, P.; Van der Kooij, A. (2007b). "Stability of Nonlinear Principal Components Analysis: An Empirical Study Using Balanced Bootstrap". *Psychological Methods*, 12(3), pp 359-379.