Statistica & Applicazioni - Vol. V, n. 1, 2007, pp. 53-77

MODELS FOR CATEGORICAL DATA: A COMPARISON BETWEEN THE RASCH MODEL AND NONLINEAR PRINCIPAL COMPONENT ANALYSIS

Eugenio Brentari*
Silvia Golia*
Marica Manisera*

SUMMARY

The paper compares two models to construct measures from the responses on a set of categorical variables, the Rasch Model and the Nonlinear (Categorical) Principal Component Analysis, and can be considered as a part of the literature about the choice between stochastic and algorithmic models. The aim is to discuss the Rasch Model and Nonlinear PCA differences and similarities, emphasizing the information that can be drawn from the data, and to compare the resulting measures.

Keywords: *Latent trait measure, Rasch model, Nonlinear* (*Categorical*) *Principal Component Analysis, simulation of multidimensional data, job satisfaction.*

1. INTRODUCTION

The construction of indicators of subjective attitudes is often one of the main goals in the economic, social and behavioural sciences. It requires statistical instruments able to measure complex concepts not directly observed, that is latent traits. To be more precise, a latent trait refers to a latent continuum or a dimension which all individuals are mapped on, based on their pattern of responses on a set of categorical variables. These categorical variables result from the submission of questionnaires with items referring to the different aspects of the concept being measured. Responses usually indicate the degree of agreement with each statement, with higher scores reflecting greater agreement. A very simple tool to assess subjective attitudes is the summated rating scale, also referred to as *raw score*. Although it is widely used, it is not appropriate in the presence of categorical variables because it assumes numerical variables linearly related to each other. Moreover, it has little inferential value, is not an interval (or ratio) measure and is affected by missing values. Therefore, raw score can only be an indication of a possible measure of the latent trait. It clearly appears the necessity to identify models and techniques able to produce quantitative measures.

* Dipartimento Metodi Quantitativi – Università degli Studi di Brescia – C.da S. Chiara, 50, 25122 BRESCIA (e-mail: {brentari;golia;manisera}@eco.unibs.it).
Despite being the result of a joint work, section 2 and sub-sections 3.2 and 4.2 should be attributed to S. Golia, section 5 and sub-sections 3.1 and 4.1 to M. Manisera. Sections 1 and 6 should be attributed to all the authors.

The two techniques compared in this study, the R*asch* M*odel* (RM) and N*on-linear* (*Categorical*) P*rincipal* C*omponent* A*nalysis* (NPCA), share the same objective, that is to give a measure of the latent trait underlying a multiple-item scale. The main difference between RM and NPCA is that RM is a *stochastic model* and NPCA is an *algorithmic model*. From the *stochastic* point of view, observations come from a data generating process with known probabilistic structure and unknown parameter values and the aim of the model is to use data to obtain the best structure of the model. On the contrary, when considering *algorithmic models*, the data generating process has unknown structure, and the aim is to use the model to identify the best representation of the data. Therefore, the present work can be considered as a part of the much-discussed question regarding the choice between stochastic and algorithmic models.

The aim is to discuss RM and NPCA differences and similarities, emphasizing the information that can be drawn from the data and to compare the measures resulting from the application of the two different statistical approaches to the data[1]. The term *measure* is here used in a broad sense: we mean a quantitative indicator that is associated to each subject involved in the study and summarizes his/her responses to a multiple-item questionnaire. Such indicator gives the evaluation of the complex concept (latent trait) under study for the considered subject and it is useful for many purposes: for example, it can be used to simply rank the individuals or as a response or explanatory variable in statistical models conceived for continuous variables (Bartholomew, 1996, p. 17).

The paper is organized as follows. Section 2 contains a brief review of the RM and NPCA. In section 3, methods to simulate multivariate data with one or two underlying latent traits are described, while section 4 reports the main results of the application of RM and NPCA to data simulated by such methods. In section 5 a real data set is analysed by RM and NPCA and the comparison between the obtained measures is discussed. Conclusions follow in section 6.

2. THE RASCH MODEL AND NONLINEAR PCA

In 1960, Georg Rasch proposed a model for dichotomously scored responses that would transform ordinal raw scores into interval scale measures when the data met the model. Since then, that model has been extended to polytomously scored items and item responses mediated by raters. Presently, the term *Rasch model* refers to a specific family of measurement models which converts raw scores into linear and reproducible measurement. Its distinguishing characteristics are: separable person and item parameters, sufficient statistics for the parameters and conjoint additivity. These features enable *specifically objective* comparisons of persons and items and allow each set of model parameters to be conditioned out of the estimation procedure for the others.

[1] A comparison between NPCA and RM was proposed by Ferrari, Annoni, Salini (2005).

RM requires unidimensionality (all the items forming the questionnaire measure only one single construct i.e. the latent trait under study) and local independence (conditionally to the latent trait, the response to a given item is independent from the responses to the other items in the questionnaire).

If the data fit the model, then the measures are objective and expressed in logit (logarithm of odds) with the attracting property of maintaining the same size (i.e. interval) over the entire continuum.

According to RM, the probability that a person answers in a given way to an item depends on the *person's ability* β (or level of latent trait) and the *item's difficulty* δ. For polytomously scored items, it can be defined by the *Rating Scale* (RSM) (Andrich, 1978) or the *Partial Credit* (PCM) (Masters, 1982) models. Therefore, assuming that there are the same $k+1$ possible ordered response categories for each item, coded as $x = 0, 1, ..., k$, the probability of a specific answer x to the item i by the subject s, following the RSM, is given by:

$$P(X_{si} = x) = \frac{\exp\left\{x(\beta_s - \delta_i) - \sum_{j=0}^{x} \tau_j\right\}}{\sum_{c=0}^{k} \exp\left\{c(\beta_s - \delta_i) - \sum_{j=0}^{c} \tau_j\right\}}, \tag{1}$$

where τ_j, called *threshold*, is the point of equal probability of categories $j-1$ and j and $\tau_0 \equiv 0$. Moreover, thresholds add up to zero.

One of the algorithmic models for measuring latent traits is Principal Component Analysis (PCA) with Optimal Scaling (Gifi, 1990; Meulman, van der Kooij, and Heiser, 2004), also known as Nonlinear PCA or Categorical PCA (CatPCA; SPSS, 2003). NPCA aims at the same goals of traditional PCA, but it is suited for variables of mixed measurement level that may not be linearly related to each other (Linting, Meulman, Groenen, and Van der Kooij, 2007).

NPCA is a descriptive approach that simultaneously reduces the dimensionality of the variables and turns the original categorical variables into quantitative variables, assigning them numerical values by a process called optimal scaling. It is possible to quantify the variables by choosing different transformations (or *scaling levels*)[2] (Michailidis and de Leeuw, 1998; Manisera, 2005). The optimally scaled values are assigned to the categories of each variable on the basis of the optimizing criterion of

[2] The optimal scaling of categorical (or numeric) data can be handled by means of nonmonotonic and monotonic transformations. In this study, only the monotonic *I*-splines transformations (i.e., *spline ordinal* scaling level) will be used (Ramsay, 1988). They preserve grouping and ordering information of the original variables. This choice is consistent with the proposed simulation study, concerning categorical ordered variables with monotonic relationships, and is in line with the RM which requires ordered categories. In the Rasch analysis categories, as measurement, divide the latent unidimensional continuum into adjacent intervals. Successive categories reflect more about the latent trait and this implies that they have to be ordered.

the procedure in use. NPCA aims at minimizing the following loss function (Gifi, 1990) simultaneously over $\boldsymbol{O}$ and the $\boldsymbol{Y}_i$'s:

$$L(\boldsymbol{O}, \boldsymbol{Y}) = m^{-1} \sum_{i=1}^{m} tr(\boldsymbol{O} - \boldsymbol{G}_i \boldsymbol{Y}_i)'(\boldsymbol{O} - \boldsymbol{G}_i \boldsymbol{Y}_i) = m^{-1} \sum_{i=1}^{m} \|\boldsymbol{O} - \boldsymbol{G}_i \boldsymbol{Y}_i\|^2 \qquad (2)$$

where $\boldsymbol{G}_i$ is the indicator matrix of item i, $i = 1, .., m$, $\boldsymbol{O}$ is the $n \times p$ matrix of *object scores* (the p-dimensional quantitative measure) for the n subjects, and $\boldsymbol{Y}_i$ is the matrix containing the category quantifications of item i. (2) evaluates the goodness of fit of the optimal quantifications $\boldsymbol{Y}_i$ for the latent variable(s) underlying the concept of interest.

The optimization problem is restricted, because orthonormalization ($\boldsymbol{1}'\boldsymbol{O} = \boldsymbol{0}'$ and $\boldsymbol{O}'\boldsymbol{O} = n\boldsymbol{I}$) and rank-one constraints ($\boldsymbol{Y}_i = \boldsymbol{y}_i\boldsymbol{a}_i'$) are imposed. In NPCA the transformed variables $\boldsymbol{q}_i = \boldsymbol{G}_i\boldsymbol{y}_i$ are usually standardized ($\boldsymbol{1}'\boldsymbol{q}_i = 0$ and $\boldsymbol{q}_i'\boldsymbol{q}_i = \boldsymbol{y}_i'\boldsymbol{G}_i'\boldsymbol{G}_i\boldsymbol{y}_i = n$). The restricted minimization of loss function (2) is derived by means of an *Alternating Least Squares* (ALS) algorithm (Gifi, 1990). A p-dimensional ($p \geq 1$) quantitative measure $\boldsymbol{O}$ of the latent trait can be created, starting from the data matrix without verifying any hypothesis (directly from the algorithmic nature of the model). It is obtained as $\sum_i \boldsymbol{q}_i\boldsymbol{a}_i'$, that is the weighted sum of the transformed variables with loadings $\boldsymbol{a}_i$ ($p \times 1$ vector) as weights. It takes into account the possible multidimensionality of the latent trait, the categorical nature of variables, and their importance in determining the measure.

From a theoretical point of view, the comparison between RM and NPCA focuses on the types of information produced by the two techniques. Being aware of this consideration helps the researcher in choosing the best approach to achieve her/his objective.

RM gives, for each item, a linear and objective measure of the difficulty to endorse that item, allowing to rank the aspects (items) from the least to the most difficult. Moreover, it provides a linear and objective measure of the level of latent trait, or ability, for each subject.

It is possible to evaluate the coherence of each single item to the latent trait and to identify anomalous subjects using diagnostic tools based on the residuals (difference between the observed and expected answer to an item by a subject) for both items and subjects (Wright and Masters, 1982). Each residual shows a piece of information about the quality of the data and the corresponding validity of the measurement model. Square standardized residuals are the ingredients of the *Outfit* and *Infit Mean Square* statistics, which help to identify problematic items and represent a tool to validate the questionnaire. Their expected value is 1 and they take values from 0 to infinity. Values close to 1 indicate little distortion of the measurement system; values lower than 1 indicate too predictable observations, while values greater than 1 indicate unpredictability and un-modelled noise.

NPCA allows the identification of the latent trait's dimensionality, that is given by the optimal number of dimensions (or components) to be maintained in the solution. Comparisons of solutions with different dimensionalities (and scaling levels) are based on goodness-of-fit indices. The eigenvalue λ_r associated with the r-th dimension gives the *Variance-Accounted-For* (*VAF*) *per dimension*. The sum λ of the first p eigenvalues λ_r is the overall goodness-of-fit index of the p-dimension solution and is equal to the *Total Variance Accounted For* (*TVAF*) in the transformed variables. NPCA assigns a quantification (that varies according to the specified scaling level) to each category of each original variable such that in the overall analysis as much as possible of the total variance of the quantified variables is accounted for. Dividing the *TVAF* by the total variance (m) leads to the *Percentage of TVAF* (*PVAF*). Another goodness-of-fit index is the *generalized Cronbach's* α (Heiser and Meulman, 1994), which coincides with Cronbach's α when $p = 1$.

The structure of the latent trait(s) can be studied by analysing the contribution of each item, given by the *loading* a_{ir}, which indicates the correlation between the i-th transformed variable and the r-th principal component, and the *PVAF per variable*.

The quantified categories give information about ordering and distance between the original categories and suggest possible recoding (Carpita and Manisera, 2006). The transformed data matrix can be used for further analyses, e.g. to rotate the NPCA solution.

3. SIMULATION METHODS

In the present study two simulation methods are considered. The first one is founded on the analytical formulation of the RM; the other originates in a data analysis context and is based on the simulation of the correlation matrix among variables.

Both methods are used to generate multivariate data with one or two latent traits in order to study the relationship between the measures obtained applying RM and NPCA.

3.1 *Simulation of data with one latent trait*

The first method employed to simulate a unique latent trait makes use of the parametric formulation of the Rasch response probability (1); the response given by the subject s to the item i is obtained as follows. First, for all the categories, the response probabilities and their cumulative sum are computed. Then, a random number rn is chosen from a uniform distribution on the interval [0,1] and compared with the cumulative sum; the first category with cumulative sum larger than rn is assigned to the response. In the simulations, the subject ability, i.e. the subject's level of latent trait, is drawn from a standardized normal distribution, hence it is possible to compare the estimated ability with the real one.

The second method is based on a data analysis approach: data sets are constructed so that a certain structure of the latent trait holds and a model error is added. The construction of such data consists of two steps: the first one is based on the algorithm by Lin and Bendel (1985) that generates random correlation matrices for a given eigenvalue structure. In this sense, the structure of the latent trait is defined by a pre-specified correlational structure: a one-dimensional latent variable underlies the data set defined by m continuous variables, based on the multinormal distribution. In much detail, the correlation matrix $\boldsymbol{C}$ between the m variables (items) correspond to a strong or a moderate one-dimensional eigenvalue structure λ (with λ indicating the dominant first eigenvalue of $\boldsymbol{C}$). It should be noted that a necessary condition to generate one-dimensional latent trait is that all the correlations between variables are equal. This has to guarantee a dominant first eigenvalue of the correlation matrix (Meulman, 1982, p. 64). Once generated $\boldsymbol{C}$, the $n \times m$ data matrix $\boldsymbol{X}$ is constructed as $\boldsymbol{X} = \boldsymbol{BS}$, where $\boldsymbol{B}$ approximates an $n \times m$ orthonormal matrix (it is randomly drawn from a multinormal distribution with zero mean and variance $1/n^2$, such that $\boldsymbol{B}'\boldsymbol{B} \approx \boldsymbol{I}$) and the $m \times m$ matrix $\boldsymbol{S}$ is given by $\boldsymbol{S} = \boldsymbol{QL}$ where $\boldsymbol{Q}$ and $\boldsymbol{L}$ result from the eigenvalue decomposition $\boldsymbol{C} = \boldsymbol{QL}^2\boldsymbol{Q}'$. The data matrix $\boldsymbol{X}$ reflects sampling variation and $\boldsymbol{X}'\boldsymbol{X}$ asymptotically equals $\boldsymbol{C}$. In order to reproduce the ordinal variables usually originated by questionnaires and used to construct measures or indicators, in the second step the continuous variables are discretized, by mapping continuous intervals into points with step functions. The number of intervals and then the number of ordered categories can be fixed at k or at k_i (a different number for each variable). Fifteen different types of discretization with $k = 5$ were chosen, starting from the proposal by van Rijckevorsel, Bettonvil and de Leeuw (1985). All the considered discretization types are the result of monotonic transformations. The frequencies corresponding to the categories of each item for the fifteen types of discretization are displayed in Table 1.

TABLE 1. - *Frequencies associated to the five categories of the items according to the fifteen types of discretization*

Discretization type	1	2	3	4	5
I	0.11	0.24	0.30	0.24	0.11
II	0.35	0.25	0.20	0.15	0.05
III	0.05	0.15	0.20	0.25	0.35
IV	0.05	0.15	0.25	0.30	0.25
V	0.07	0.08	0.22	0.28	0.35
VI	0.25	0.30	0.25	0.15	0.05
VII	0.15	0.30	0.30	0.15	0.10
VIII	0.15	0.30	0.30	0.15	0.10
IX	0.18	0.32	0.25	0.15	0.10
X	0.10	0.15	0.30	0.30	0.15
XI	0.10	0.15	0.25	0.33	0.17
XII	0.45	0.25	0.15	0.10	0.05
XIII	0.05	0.10	0.15	0.25	0.45
XIV	0.05	0.10	0.15	0.30	0.40
XV	0.40	0.30	0.15	0.10	0.05

In order to create variables which correspond to the responses given to a questionnaire without any reversed item, a one-dimensional PCA on the continuous variables is performed, in order to identify which variables have a negative loading on the found component. The identified variables refer to the reversed items in the questionnaire and have to be recoded so that all the correlations in $\boldsymbol{C}$ have the same sign.

RM needs items with increasing ordinal categories since categories are considered as measurement in the sense that they divide the latent unidimensional continuum into adjacent intervals and successive categories reflect successively more about the latent trait. The Rasch analysis of data generated without the previous adjustment highlights the presence of reversed items if the point-biserial correlation coefficient, which measures the correlation between each item response score and the total person test score, assumes a negative value. NPCA, on the contrary, does not require the adjustment of reversed items because it constructs indicators by giving an opposite loading to the reversed items (for example, if the unreversed items have positive loadings, the loadings of the reversed items are negative).

Different types of discretization have been created in order to simulate items with different *mean difficulty*, using a RM terminology. In fact, whilst the application of NPCA to data sets created with only three discretization types does not reveal any problem with the questionnaire, the item person map (Figure 1), that is the map that displays simultaneously the distribution of subject ability (on the left side of the vertical line) and item difficulties (on the right side of the vertical line), obtained by RM, shows that the items are not well distributed along the latent continuum.

Using Rasch analysis as a tool to validate the questionnaire, most of the items should be replaced by others with different difficulty.

Data sets with items better distributed along the item person map can be obtained by applying a number of different discretization types to the continuous variables, leading to different frequency distribution for the items. In this sense, the application

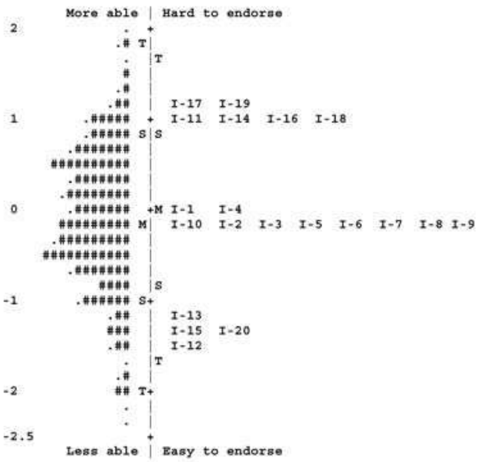


FIGURE 1. - *Item person map for data with only three types of discretization*

of the RM allowed the improvement of the second simulation method in order to create data sets corresponding to more realistic questionnaires.

3.2 *Simulation of data with two latent traits*

The first method used to simulate data with two latent traits consists in identifying two groups of items which require two different abilities and then in applying the response probability (1), according to the scheme described in 3.1. Hence, following Smith (2002), two independent standard normal ability distributions (Z and W) are generated. From these two distributions, a third one is obtained as:

$$W_c = \left(h \cdot Z + \sqrt{1-h} \cdot W\right) \cdot S_c + M_c, \quad (3)$$

where S_c and M_c are the standard deviation and the mean of the W_c distribution, respectively. The W_c distribution represents the ability distribution correlated with Z at a specific level of h ($h \in [0, 1]$). Low values of h indicate a low degree of correlation between Z and W_c so that the two groups of items require two abilities which are sparely related. On the contrary, high values of h indicate a high correlation between Z and W_c so that the two groups of items require two abilities highly related and the multidimensionality is lighter.

The second method is a generalization of the analogous method described in section 3.1 for the one-dimensional case. The structure of the two latent traits is again defined by a pre-specified correlational structure; the only difference concerns the definition of the correlation matrix: $\boldsymbol{C}$ is now a block-diagonal matrix where the first block in the diagonal, $\boldsymbol{C}_{\mathrm{I}}$, contains the correlations between m_{I} variables and the second block in the diagonal, $\boldsymbol{C}_{\mathrm{II}}$, contains the correlations between m_{II} variables (obviously, $m_{\mathrm{I}} + m_{\mathrm{II}} = m$). The off-diagonal blocks consist of zeroes, such that the two groups of variables do not correlate with each other. Both groups of variables correspond to a strong or a moderate one-dimensional eigenvalue structure λ_i (λ_i with $i = \mathrm{I}, \mathrm{II}$ indicating the dominant first eigenvalue of block C_i corresponding to the i-th group of variables). In this way, two latent variables, based on the multinormal distribution, underlie the data set, with m_{I} continuous variables forming the first latent trait and m_{II} continuous variables the second latent trait. Checking for the reversed items is done here by means of a two-dimensional PCA on the continuous variables. The continuous data matrix is then discretized as explained is section 3.1 and any reversed item is adjusted.

4. SIMULATION RESULTS

In the present section the results of the analysis of simulated multivariate data sets with one or two latent traits are presented. A 5-level Likert response scale for each item and a test size of $m = 20$ items have been chosen. Each data set reports the responses of $n = 500$ subjects. The Rasch analysis and NPCA were performed using Winsteps 3.57 (Linacre, 2005) and SPSS 13.0 respectively.

4.1 *Data sets with one latent trait*

The first data set to be analysed (*data1*) is generated using the response probability (1) with items' mean difficulties set equal to [-1.81, -1.73, -1.59, -1.48, -1.36, -0.79, -0.64, -0.44, -0.25, -0.08, 0.07, 0.19, 0.44, 0.76, 0.91, 1.24, 1.38, 1.52, 1.73, 1.92] and the 4 threshold parameters equal to [-0.9, -0.4, 0.4, 0.9].

The first step in the Rasch analysis consists in verifying the assumptions of unidimensionality and local independence.

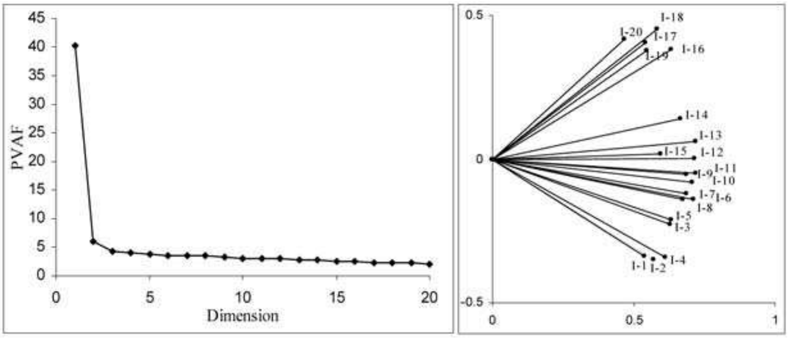
In order to verify the first assumption, the *outfit* and *infit mean square* statistics and PCA on standardized residuals are evaluated. The purpose of the un-rotated PCA on standardized residuals used in the Rasch context is to not find shared factors. The underlying hypothesis is that there is only one dimension, called the *Rasch dimension*, captured by the model so that the residuals do not contain other significant dimensions. Hence, the purpose is to verify the absence of other significant dimensions. In the present data set no item misfits and the first eigenvalue $(\lambda_1 = 1.2)$ is consistent with the unidimensional assumption.

Pearson correlation is one of the tools for testing local independence. For each pair of items, the Pearson correlation is computed between residuals across all subjects who responded to both items. Potentially locally dependent pairs of items have high positive or negative correlations. In the present data set no couple of items shows high positive or negative inter-item correlation.

Moreover, the original ranking of the items is reproduced.

The application of NPCA to *data1* clearly reveals a one-dimensional latent trait. The scree test suggests the one-dimension solution, in fact the scree plot of the solution retaining all the 20 dimensions (Figure 2a) shows an elbow in correspondence of the second dimension, therefore the optimal number of dimensions to be maintained in the solution is one.

The *eigenvalue greater than 1* criterion would suggest to retain two dimensions in the solution, because the first two eigenvalues of the 20-dimension solution are higher than 1 (8.04 and 1.21). However, λ_2 is much smaller than λ_1 $(\lambda_1/\lambda_2 = 6.64)$. The two criteria (scree test and eigenvalue greater than one) must be evaluated with care: because the solutions in NPCA are *non-nested* (Meulman et al., 2004), they only give indications about the number of dimensions to be retained in the solution. Moreover, the two-dimension solution makes it clear that it is appropriate to choose the one-dimension solution for at least two reasons: (i) the improvement in the goodness of fit obtained going from the one- to the two-dimension solution is not remarkable (PVAF increases from 41.26 to 47.15 and Cronbach's α from 0.92 to 0.94); (ii) even if there is a group of five variables (I-16:I-20) with a non negligible loading on the second dimension, all the 20 variables contribute more to the first dimension than to the second one (Figure 2b). The quantified categories of the 20 items show that



(a) (b)

FIGURE 2. - (a) *Scree plot of the NPCA 20-dimension solution and* (b) *loadings of the NPCA 2-dimension solution for data1*

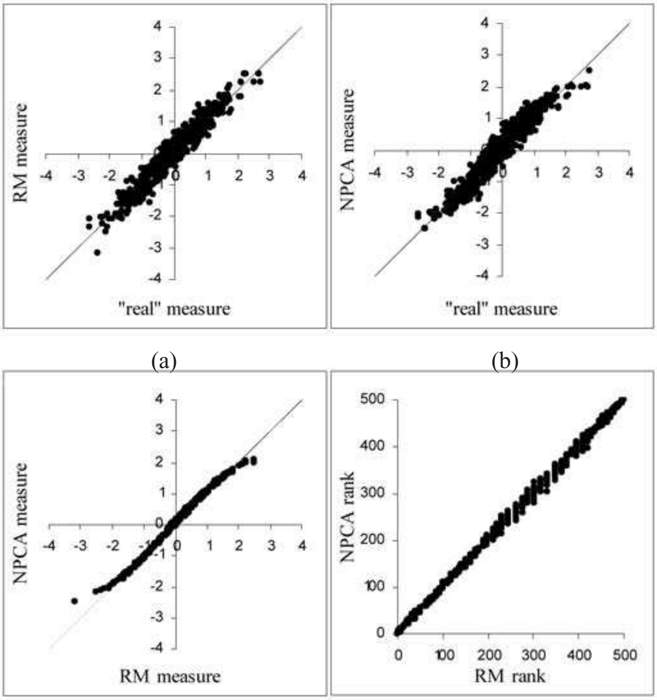
the original categories are ordered but not always equally spaced, indicating that the choice of the spline ordinal scaling level is appropriate. There is no need to recode variables by merging categories.

Figure 3a shows the comparison between the standardized real (simulated) level of latent trait and the standardized Rasch measure. The good agreement between the two measures is confirmed by the high value of the Pearson correlation ρ (0.963). Similar conclusions can be drawn by Figure 3b, where the standardized real level of latent trait is plotted against the NPCA measure derived from the one-dimension solution ($\rho = 0.961$). The standardized Rasch and NPCA measures are highly correlated ($\rho = 0.991$), as shown in Figure 3c, and the subjects ranking based on the two measures is substantially the same (Figure 3d): the Spearman correlation coefficient equals 0.999.

Using the second simulation method we generated two data sets[3] with different values of λ/m: *data2* with $\lambda/m = 0.8$ and *data3* with $\lambda/m = 0.5$. In both cases, the continuous variables were discretized according to the following scheme: each type of discretization (see Table 1) was used to discretize one single variable, except for types I, II, III used to discretize 2, 3, and 3 variables respectively.

In both cases, the un-rotated PCA on standardized residuals used in the Rasch context supplies a first eigenvalue consistent with the unidimensionality assumption ($\lambda_1 = 1.5$ for *data2* and $\lambda_1 = 1.3$ for *data3*) and no couple of items shows high positive or negative inter-item Pearson correlation. In the case of $(\lambda/m) = 0.8$, few items misfit whereas in case of $\lambda/m = 0.5$ no item misfits.

[3] The correlations computed among the continuous variables of *data2* (before the discretization step) range between 0.74 and 0.81. The dominant eigenvalue equals 15.78, the subsequent 19 eigenvalues vary between 0.15 and 0.31. In *data3*, the correlations among the continuous variables range from 0.35 and 0.50. The dominant eigenvalue equals 9.15, the others vary between 0.37 and 0.76. Although equality of correlations is a necessary condition for the existence of a one-dimensional latent trait, the correlations in each data set are not perfectly equal due to the effect of sampling. Clearly, the NPCA results are also affected by the effects of discretization.



(c) (d)

FIGURE 3. - *Scatterplots of* (a) *simulated level of latent trait and RM measure* ($\rho = 0.963$) (b) *simulated level of latent trait and NPCA measure* ($\rho = 0.961$) (c) *RM and NPCA measures* ($\rho = 0.991$) (d) *RM and NPCA ranks of subjects* (*Spearman* $\rho = 0.999$)

The NPCA 20-dimension solutions of both data sets show a scree plot confirming the one-dimensional latent trait (Figure 4). The one-dimension solution shows a *PVAF* equal to 73.14 ($\alpha = 0.98$) and 40.52 ($\alpha = 0.92$) for *data2* and *data3* respectively.

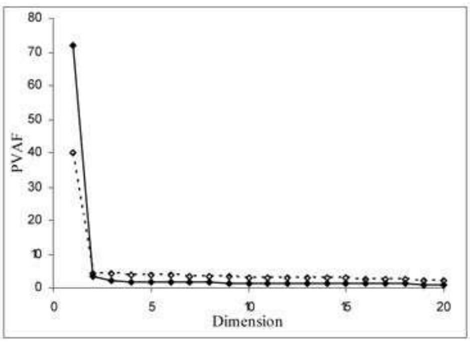
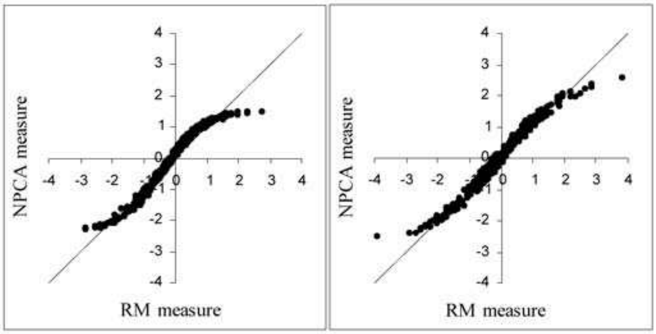


FIGURE 4. - *Scree plot of the 20-dimension NPCA solutions for data2* (*black line*) *and data3* (*dashed line*)



(a) (b)

FIGURE 5. - *Scatterplots of RM and NPCA measures for* (a) *data2* ($\rho = 0.976$) *and* (b) *data3* ($\rho = 0.981$)

Here and in the following applications the *PVAF*'s found by NPCA are always smaller than the simulated ones due to sampling effects but also because the discretization procedure leads to a loss of information in the data: the discretized data are certainly less informative than the continuous ones and this is reflected in the percentage of *VAF* of whatever NPCA solution.

The NPCA quantified categories of the 20 items show that the original categories are ordered and nearly always equally spaced. There is no need to recode variables by merging categories.

In Figure 5 the standardized Rasch measure is plotted against the NPCA measure for both *data2* and *data3*. The good agreement between the measures is confirmed by high values of ρ (0.976 and 0.981 for *data2* and *data3* respectively). The performance of the two methods in measuring extreme subjects differs especially when the latent trait is stronger, as evident in Figure 5.

If the strength of the latent trait is moderate (*data3*), the second simulation method supplies data comparable to the ones given by the method based on RM; the *PVAF* is 41.26 for *data1* and 40.52 for *data3*.

4.2 *Data sets with two latent traits*

Following the response probability (1), the normal ability distribution X is used to generate the responses to 15 items with mean difficulties [-1.81, -1.73, -1.59, -1.36, -0.79, -0.64, -0.25, -0.08, 0.07, 0.44, 0.76, 1.24, 1.38, 1.52, 1.92] whereas the ability distribution W_c, obtained applying (3) with $M_c = 0$, $S_c = 1$ *and* $h = 0.1$ (*data4*) and $h = 0.6$ (*data5*), is used to generate the responses to the last 5 items with mean difficulty [-1.48, -0.44, 0.19, 0.91, 1.73]. The 4 threshold parameters are set equal to [-0.9, -0.4, 0.4, 0.9].

In both cases, the first eigenvalue ($\lambda_1 = 3.9$ if $h = 0.1$ and $\lambda_1 = 2.8$ if $h = 0.6$), obtained applying the un-rotated PCA on the Rasch standardized residuals, suggests

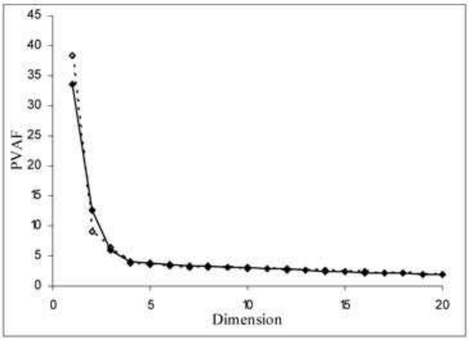
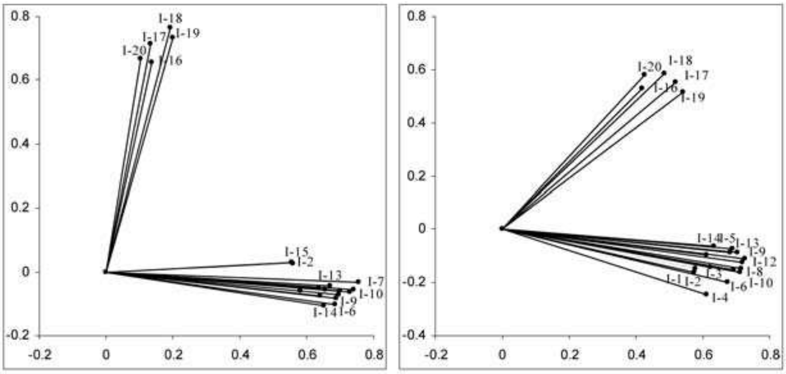


FIGURE 6. - *Scree plot of the 20-dimension NPCA solutions for data4* (*black line*) *and data5* (*dashed line*)

the presence of a second latent construct in the data. Only the five items responsible for the second construct misfit and the couples obtained from their combinations show fairly high positive inter-item Pearson correlation.

The 20-dimension NPCA solution suggests that the latent trait underlying both data sets is certainly not one-dimensional, even if in both cases the first dimension catches the majority of information in the original variables (this is a characteristic of NPCA). The scree plot of the 20-dimension solutions (Figure 6) shows the opportunity to consider more than one dimension for both data sets. In both *data4* and *data5*, the first three eigenvalues of the solution are greater than 1 (6.6, 2.5, 1.2 and 7.6, 1.8, 1.3 for the *data4* and *data5* respectively) and the eigenvalues of the three-dimension solution are still greater than one (6.6, 2.6, 1.5 and 7.7, 1.8 and 1.4 for the two data sets respectively), suggesting a three-dimension solution. But the inspection of loadings of the three-dimension solution suggests to retain two dimensions, because no variables load higher on the third dimension than on one of the other two dimensions.

In *data4*, the two-dimension solution shows a *PVAF* equal to 46.79 (33.92 and 12.87 for the two dimensions separately) and a generalized Cronbach's α of 0.94. The loadings (Figure 7a) indicate that the first dimension is composed of the first 15 items (with loadings varying between 0.56 and 0.75), while the second dimension is composed by the items I-16:I-20 (with loadings varying between 0.14 and 0.23), perfectly in line with the simulated abilities, differing between the first 15 items and the last 5 items. The separation between the two groups of items suggests the possible existence of two distinct latent traits. In *data5*, the two-dimension solution shows a *PVAF* equal to 48.05 (38.88 and 9.17 for the two dimensions separately) and a generalized Cronbach's α of 0.94. Also in this case the loadings (Figure 7b) separate the first 15 items, loading higher on the first dimension (loadings range from 0.57 to 0.73), from the last 5 items, loading high on the second dimension (loadings range from 0.43 to 0.54). The separation between the two groups of items in Figure 7b is less evident than in Figure 7a, nevertheless it suggests the possible existence of two distinct latent traits.



(a) (b)

FIGURE 7. - *Loadings of the 2-dimension NPCA solutions for* (a) *data4 and* (b) *data5*

We can explain the differences in the *PVAF*'s per dimension and in the loadings for *data4* and *data*5 in the following way. When $h = 0.6$ the 20 items (and, therefore, the two latent traits) share a non negligible part of information, which is captured by the first dimension (also evident in Figure 7b). On the contrary, when $h = 0.1$, the 20 items (and, therefore, the two latent traits) share less information: the second dimension becomes more informative and, then, it is more important to be retained in the final solution (also evident in Figure 7a). Moreover, it is interesting to see that the two groups of items are much closer in Figure 7b than in 7a: this reflects the different correlation level between the two latent traits underlying the data.

The NPCA one-dimension solutions of *data4* and *data5*, supplying the one-dimensional measures to be compared with the RM measures, explain 34.10% ($\alpha = 0.90$) and 38.99% ($\alpha = 0.92$) of *VAF* respectively. As expected, NPCA goodness-of-fit indices of the one-dimension solution allow us to appreciate that the first latent construct is stronger in *data5* than in *data4*.

The comparison between Rasch and NPCA measures can only be done if the data come from a unidimensional questionnaire (that is data related to only one latent trait). When the data reveal the presence of more than one latent trait, the Rasch measure must be computed on each sub-group of items composing a separate latent trait. In the present case, the Rasch measure estimated using the 15 items which do not misfit is used for the comparison with the NPCA measure derived from the one-dimension solution[4].

In Figure 8 the standardized Rasch measure is plotted against the NPCA measure for *data4* and *data5*. In both cases there is a good agreement between measures

[4] In the presence of two latent traits, a one-dimensional NPCA measure can be obtained by (i) the one-dimension solution, (ii) the first dimension of the two-dimension solution, or (iii) the one-dimension solution considering only the sub-group of items loading high on the first dimension. The three NPCA measures are comparable (highly correlated) because all of them take essentially into account only that sub-group of items. For example, the correlations among the three measures obtained for *data6* vary from 0.996 to 1.

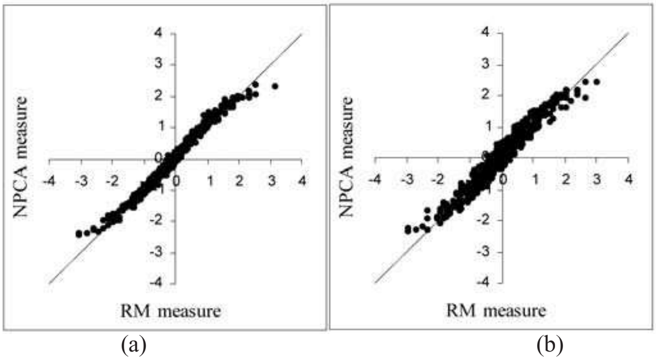


FIGURE 8. - *Scatterplots of RM and NPCA measures for* (a) $h = 0.1$ ($\rho = 0.989$) *and* (b) $h = 0.6$ ($\rho = 0.971$)

which is also highlighted by the high values of ρ (0.989 and 0.971 respectively). It can be noted that a stronger accordance between measures shows up when $h = 0.1$, that is when the two latent trains are more separate. It must be reminded that the Rasch measure is computed taking into account only the first 15 items and the NPCA measure is computed as weighted sum of the transformed variables with weights equal to loadings. The loadings associated to items I-16:I-20 vary in (0.14, 0.23) for *data4* and in (0.43, 0.54) for *data5*. Therefore, when $h = 0.1$, the impact of the last 5 items (forming the second trait) on the construction of the one-dimensional NPCA measure is lower.

Other two data sets with an underlying two-dimensional latent trait were simulated according to the second method described in 3.2. To generate the first data set (*data6*), we set $m_{\mathrm{I}} = 15$ with $\lambda_{\mathrm{I}}/m_{\mathrm{I}} = 0.8$ and $m_{\mathrm{II}} = 5$ with $\lambda_{\mathrm{II}}/m_{\mathrm{II}} = 0.5$. The second data set (*data7*) was obtained with $m_{\mathrm{I}} = 15$, $\lambda_{\mathrm{I}}/m_{\mathrm{I}} = 0.5$, $m_{\mathrm{II}} = 5$, and $\lambda_2/m_{\mathrm{II}} = 0.5$. The obtained continuous variables were discretized according to the same scheme described for *data2* and *data3*.

In both cases, the first eigenvalue ($\lambda_1 = 6.3$ for *data6* and $\lambda_1 = 3.6$ for *data7*), obtained applying the un-rotated PCA on the Rasch standardized residuals, suggests the presence of a second latent construct in the data. Moreover, if *data6* is considered, then all the 20 items misfit and most of the couples of items show fairly high positive or negative inter-item Pearson correlation, whereas if *data7* is analysed, then most of the 20 items misfit and the couples obtained combining the five items responsible for the second construct show fairly high positive inter-item Pearson correlation.

The application of NPCA to the two data sets shows the presence of two latent traits in a clearer manner with respect to the results obtained for *data4* and *data5*. In fact, the 20-dimension solution shows only two eigenvalues greater than one: 10.75 and 2.42 for *data6* and 6.74 and 2.28 for *data7*; these *VAF*'s are smaller than the simulated ones ($\lambda_{\mathrm{I}} = 12$ and $\lambda_{\mathrm{II}} = 2.5$ for *data6* and $\lambda_{\mathrm{I}} = 7.5$ and $\lambda_{\mathrm{II}} = 2.5$ for *data7*) due to both sampling and discretization effects. The scree plot of the 20-dimension solutions obtained for both data sets (Figure 9) confirms the conclusion that the latent trait underlying the data is two-dimensional.

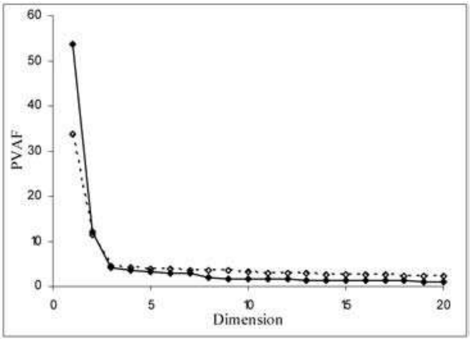
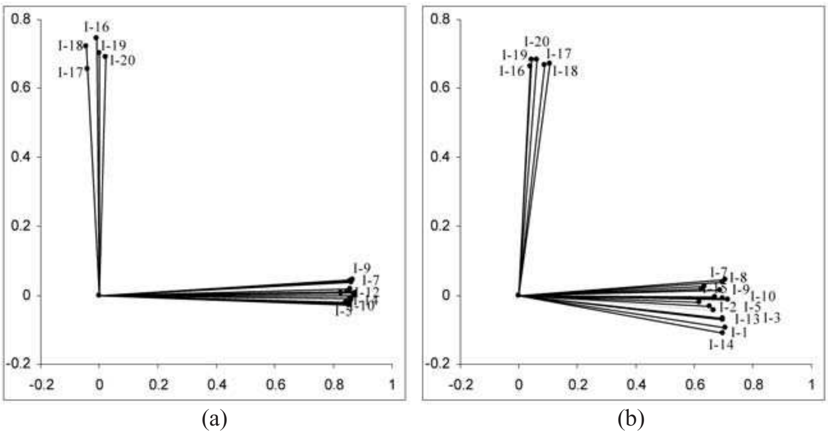


FIGURE 9. - *Scree plot of the NPCA 20-dimension solutions for data6* (*black line*) *and data7* (*dashed line*)

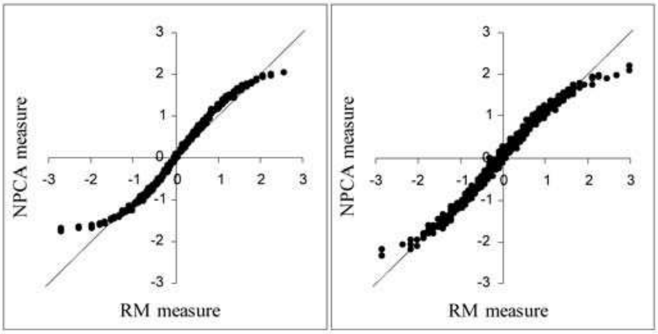
In *data6* the two-dimension solution shows a *PVAF* equal to 67.36 (54.91 and 12.45 for the two dimensions separately) and $\alpha = 0.97$. The two-dimension solution obtained for *data7* explains the 46.29% of *VAF* (34.70% and 11.59% for the two dimensions separately) and $\alpha = 0.94$.

The component loadings of the two solutions (Figure 10) well separate items I-1:I-15 (contributing to the first dimension) from items I-16:I-20 (forming the second dimension). Each dimension can be seen as a distinct latent trait. It is interesting to note that when $\lambda_{\mathrm{I}}/m_{\mathrm{I}} = 0.8$, the 15 items are represented in the plot by vectors very close to each other (Figure 10a), whereas when $\lambda_{\mathrm{I}}/m_{\mathrm{I}} = 0.5$ the same vectors are less compact (Figure 10b); this graphically represents the strength of the first latent dimension and corresponds to the larger (varying from 0.64 to 0.76) or smaller (from 0.34 to 0.51) correlations between transformed variables existing in the solutions for *data6* and *data7* respectively. The distance between groups of vectors directly derives from the simulated



(a) (b)

FIGURE 10. - *Loadings of the NPCA 2-dimension solutions for* (a) *data6 and* (b) *data7*



(a) (b)

FIGURE 11. - *Scatterplots of RM and NPCA measures for* (a) *data6* ($\rho = 0.979$) *and* (b) *data7* ($\rho = 0.982$)

uncorrelation between the two sub-groups of variables (the two blocks out of the diagonal in the correlation matrix $\boldsymbol{C}$ are composed of all zeroes).

The one-dimension NPCA solutions for *data6* and *data7* show *PVAF*'s equal to 54.96 ($\alpha = 0.96$) and 34.73 ($\alpha = 0.90$) respectively. As expected, NPCA goodness-of-fit indices of the one-dimensional solution allow us to appreciate that the first latent construct in the first case is stronger than in the second one.

In Figure 11 the standardized Rasch measure is plotted against the NPCA measure for *data6* and *data7*. In both cases there is a good agreement between measures ($\rho = 0.979$ and 0.982 respectively). It should be noted that *data6* presents a less realistic situation than *data7*: in *data6*, the strength of latent construct underlying items I-1:I-15 (and measured by *PVAF*) is notably higher than in *data7*, where it is similar to *data4* and *data5*, generated according to (1). This fact can be used to explain why Figure 11b, more than Figure 11a, is comparable with Figure 8. Moreover, the similarity is clearer when considering Figure 8a, because the separation between the two latent traits created by the second simulation method is analogous to the one obtained when $h = 0.1$.

5. REAL DATA

The available data come from the first national survey concerning the social service sector realized in Italy (Borzaga, 2000; Borzaga and Musella, 2003) requested by *Fondazione Italiana per il Volontariato* and *Fondazione Europa Occupazione*[5]. Five different types of organizations (social cooperatives, public bodies, for profit, lay non-profit and religious non-profit organizations) from 15 Italian provinces were involved in the survey. The questionnaire consisted of two parts; one refers to personal and professional information while the other to work-related attitudes, job satisfaction, relationships with end users, colleagues and superiors.

[5] The survey has been carried out by a wide group of scholars coordinated by Prof. C. Borzaga at Istituto Studi Sviluppo Aziendale Nonprofit (ISSAN), Trento University.

The present analysis refers to the data related to the satisfaction section of the questionnaire which includes the overall satisfaction as well as 14 aspects of work such as appreciation, job conditions, pay or co-workers. For each work aspect, respondents were asked to indicate how satisfied they were with it, choosing a score from 1 (very dissatisfied) to 7 (highly satisfied). Only workers employed in social cooperatives were considered so that the size of the data set is 541.

Preliminary analyses (Brentari and Golia, 2005; Manisera, Dusseldorp, and Van der Kooij, 2005) suggested to merge together the second and the third categories, obtaining a 6-level Likert response scale for each item and to leave out the first and last items so that 13 items, labelled as *Growth*, *Indep*, *Recog*, *Variety*, *Physical*, *Benefit*, *Wage*, *Hours*, *Carprom*, *Carprosp*, *Certainty*, *Superiors* and *Colleagues*, have been included in the analysis. Appendix reports the detailed description of the 13 items considered.

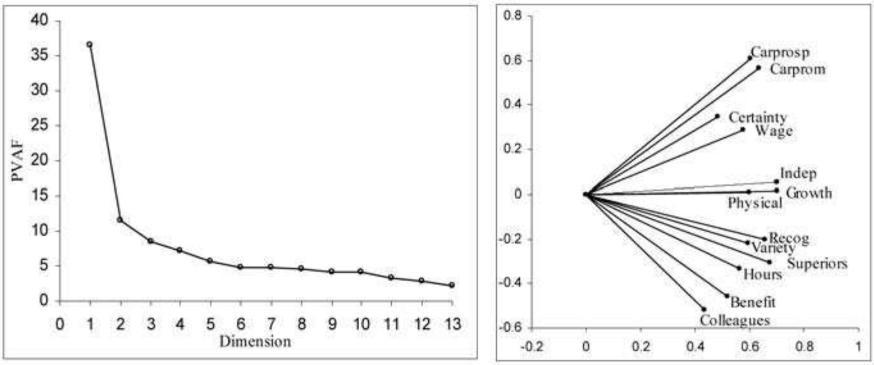
For every item the percentage of nonresponses is calculated. These percentages are small, varying between 0.2% (one worker) and 2% (ten workers), and do not affect the analysis.

The data were analysed applying the RSM (1), setting the mean of item difficulty estimates as 0.0 logits and using the (unconditional) maximum likelihood estimation method, and NPCA with spline ordinal scaling level and variable mode imputation of missing values. NPCA transformations of variables obtained without ordering restrictions show an increasing trend, suggesting that the used scaling level (requiring the ordering of the original categories) leads to the same results.

The two methods give the same indication on the number of latent traits underlying the data. The first eigenvalue ($\lambda_1 = 2.3$), obtained applying the un-rotated PCA on the Rasch standardized residuals, would suggest the presence of a second latent construct in the data. Nevertheless, the two items responsible for this second trait, *Carprom* and *Carprosp*, do not misfit ($infit = 1.02$; $outfit = 1$ and $infit = 1.03$; $outfit = 1.04$ respectively). Moreover, their inter-item correlation is rather high (0.52), suggesting that the second trait can be treated as a sub-dimension of job satisfaction (two variables measure the same aspect of job satisfaction), rather than a distinct second trait not related to the main trait. The same conclusion can be obtained by NPCA[6]. The scree plot of the 13-dimension solution (Figure 12) does not give a clear indication on the number of dimensions to be maintained in the solution, because the elbow would suggest to choose one dimension, but the *PVAF* associated to the subsequent dimensions is non negligible.

Also the *eigenvalue greater than one* criterion does not give a clear indication; a strict application of such criterion would suggest a three-dimension solution: three eigenvalues greater than one (4.7, 1.5, 1.1) are in the 13-dimension solution and so is for the three-dimension solution (4.7, 1.6, 1.1). However, the third eigenvalue is not so far from one and the interpretability of the three-dimension solution ($PVAF = 57.29$; $\alpha = 0.94$) is very difficult: two out of the three dimensions do not satisfy the Stevens'

6 The data used in the present work are a subset of those analysed in Brentari and Golia (2005) and in Manisera et al. (2005).



(a) (b)

FIGURE 12. - (a) *Scree plot of the NPCA 13-dimension solution and* (b) *loadings of the NPCA 2-dimension solution for the job satisfaction data*

criterion (Stevens, 1992), suggesting that reliable components should have a minimum of four loadings above 0.60 (in absolute value). The 2-dimension solution explains 49.01% of the *TVAF* (36.15% and 12.86% for the first and the second dimension respectively) and shows a generalized Cronbach's α of 0.91. Also the loadings of the 2-dimension solution make the interpretation of components difficult (Figure 12b). In fact, the second component does not satisfy the Stevens' criterion: only two variables, *Carprom* and *Carprosp*, load high on the second component, with a loading of 0.6 each. This is certainly due to the high correlation between those items: the Pearson correlation coefficient between the transformed variables (in the 2-dimension solution) of *Carprom* and *Carprosp* equals 0.73, while the correlations between all the other pairs of variables vary between 0.05 and 0.52). This means that *Carprom* and *Carprosp* form together a unique factor, as mentioned before.

More interpretable two or three-dimension solutions can be obtained by rotating the corresponding NPCA solutions. This facilitates the identification of job satisfaction sub-dimensions, feasible by NPCA. However, in order to obtain a one-dimensional measure comparable with the RM measure, we preferred to consider here only un-rotated solutions. In fact, in a rotated two or three-dimension solution, all the components are necessary to define the job satisfaction indicator, both because all items contribute to one of the factors and because the *VAF* is distributed among all the components. Therefore, we cannot compare only one component of a rotated solution with the RM measure. Instead, the (un-rotated) one-dimension NPCA solution is immediately comparable with the Rasch measure, because it explains the majority of *VAF* and all variables contribute to its definition.

Concluding, in this data set job satisfaction can be considered as a one-dimensional latent trait. The *PVAF* in the one-dimension solution is 37.27 and the Cronbach's α equals 0.86. The career satisfaction can be considered as a sub-dimension of job satisfaction, also fitting well in the one-dimension solution: in fact, the loadings of *Carprom* and *Carprosp* were also high (0.6 each one) on the component of the one-dimension solution.

The two methods give similar conclusions also with reference to the properties of the categories in the data. The NPCA transformations of variables without ordering restrictions show an increasing trend. Moreover, the trend obtained requiring ordering of the categories is nonlinear, confirming the choice of the NPCA over a standard or linear PCA. The trend of the transformations indicates that the original categories of the items are ordered (increasing trend) but not equally-spaced. The respondents were not always able to perfectly distinguish between different original categories. However, each category obtains its own quantification, therefore there is no need to merge categories. Note that merging categories having the same quantification leads to a solution with the same goodness of fit. Conversely, merging categories having different quantifications causes a loss in the goodness of fit. Equivalent conclusions can be drawn from the analysis of the RM thresholds τ_j which result: $\tau_1 = -1.31$,

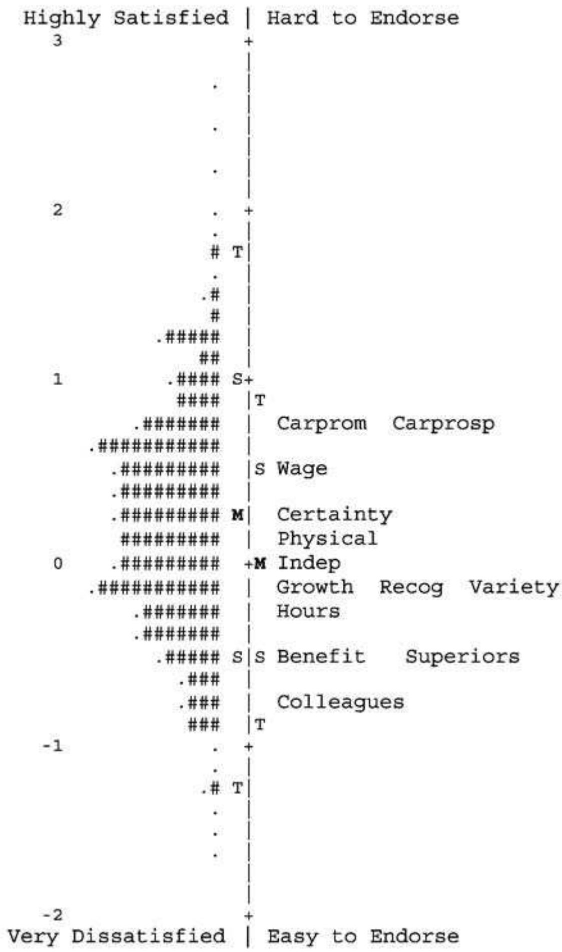


FIGURE 13. - *Map of item and worker locations (each # corresponds to 4 workers)*

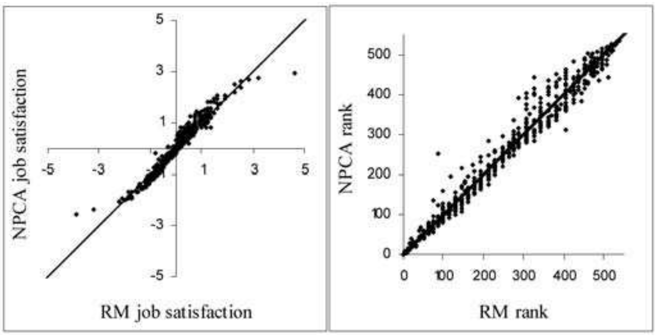
$\tau_2 = -0.28$, $\tau_3 = -0.03$, $\tau_4 = 0.39$, and $\tau_5 = 1.23$. The estimated thresholds are ordered, then the ordering is a data and not only a model property. It has to be noted that when categories and thresholds are disordered, merging categories will improve the item fit and the overall scale and may reveal the effective number and ordering of categories, nevertheless if categories and thresholds are ordered, then collapsing to a smaller number of categories will give a poorer fit of the model.

A goodness-of-fit index for RM is person reliability index which is an estimate of the replicability of people placement that can be expected if the sample was given another set of items measuring the same latent construct. It is bounded by 0 and 1 and can be computed in the presence of missing values too. In the present analysis it takes value 0.83; such value provides a good level of confidence that the worker placement would be reproducible with a different questionnaire measuring the same construct.

The item person map (Figure 13) simultaneously shows the distribution of worker satisfaction (on the left side of the vertical line) and item difficulties (on the right side of the vertical line). ‘‘M’’ marks the worker and item mean, ‘‘S’’ is one sample standard deviation away from the mean and ‘‘T’’ is two sample standard deviations away from the mean. The workers located at the upper end of the scale are the most satisfied; on the contrary, the ones located at the lower end are the least satisfied. Moreover, workers with satisfaction measure larger than 0.26 exhibit a satisfaction level above the average. The items less likely to be satisfactory are on top, so *Carprom* and *Carprosp* are the least satisfactory items followed by *Wage*. On the contrary, *Colleagues* is the most satisfactory item, followed by *Benefit* and *Superiors*.

Therefore, the information on the replicability of people placement and on the people/item ranking results from the application of RM.

In this data set, the two models can be indifferently used to obtain either the job satisfaction measure or the ranking from the least to the most satisfied worker. Figure 14a shows the comparison between the standardized Rasch and NPCA measures. Like in the analysis of simulated data with one-dimensional latent trait (sub-section 4.1), there is a good agreement between the two measures, also confirmed by $\rho = 0.969$.



(a) (b)

FIGURE 14. - *Scatterplots of* (a) *RM and NPCA job satisfaction measures* (b) *RM and NPCA ranking of workers with respect to their job satisfaction*

The ranking of workers based on the two measures is substantially the same (Figure 14b) with Spearman correlation coefficient equal to 0.986.

6. CONCLUSIONS

The paper compares RM and NPCA, emphasizing differences and similarities. The analysis of simulated data allows to compare the Rasch and the NPCA measures, the latter obtained by the one-dimension solution. A substantial concordance between the two standardized measures is pointed out.

The analysis of simulated multivariate data sets with two latent traits let us to study the impact of the bi-dimensionality on the measures. Rasch measure maintains the properties of linearity and objectivity if the unidimensionality assumption holds, so it has been calculated on the sub-group of items referred to the *dominant* latent trait, that is the one defined by higher number of items. The comparison between the two measures required a one-dimensional NPCA measure, which can be obtained in turn by (i) the one-dimension solution, (ii) the first dimension of the two-dimension solution, or (iii) the one-dimension solution only considering the sub-group of items loading high on the first dimension. Given that the three measures are usually highly correlated when the two latent traits are well separated, the one-dimension solution (i) has been considered. Again, there is a good agreement between the two measures.

When the two distinct latent traits are described by a high and equal number of items, two separate Rasch measures must be computed for each subject. Conversely, the NPCA measure can be directly obtained from the two-dimension solution and each dimension is comparable with the corresponding Rasch measure.

It should be remarked that the presence of a multidimensional latent trait could imply the existence of either more than one latent trait or only one latent trait affected by sub-dimensions. It is necessary to dispose of the detailed description of the items and a good experience in the research field in order to understand if the dimensions are separate latent traits or sub-scales of the unique latent trait.

The analysis of the real data set highlights the common information, such as unidimensionality of the latent trait, and a substantial concordance in both measures and ranking of workers.

In our opinion, the two models are not competing but completing: the choice depends on the objectives of the research, specifically the type of information to be drawn from the data. On one side, if the goal is to achieve a measure of the latent trait on n subjects (or their ranking) the two models seem to be interchangeable. On the other side, the two models can be jointly used in order to get the more information as possible from the data: for example, RM provides the items' ranking while NPCA identifies possible sub-dimensions. Moreover, the data themselves can determine what model can be used. For example, if the hypotheses underlying the RM are not satisfied by the data, such a model cannot be applied; conversely, NPCA is nearly always applicable, but in some cases the results are unstable (for example, when the marginal frequencies of some categories are very small).

Future developments involve a stability study in order to make the results of the present work more general. Such study should concern the stability of: (a) the NPCA eigenvalues, Cronbach's α, loadings, quantifications, and object scores (Gifi, 1990; Michailidis and de Leeuw, 1998), (b) the RM parameters, and consequently (c) the outcomes of the comparison between the NPCA and RM measures.

ACKNOWLEDGEMENT

The authors wish to thank M. Carpita for his useful and valuable discussions and the two anonymous reviewers.

This work is financially supported by MIUR-PRIN 2004, prot. 2004130924-004.

RIASSUNTO

Il presente lavoro mette a confronto due modelli utilizzabili per la creazione di una misura a partire da un insieme di variabili categoriali, il modello di Rasch e l'analisi delle componenti principali nonlineare, e può essere considerato parte della letteratura riguardante la scelta tra modelli stocastici ed algoritmici. L'obiettivo è quello di discutere differenze e punti di contatto dei due metodi, dando risalto alle informazioni peculiari derivanti dalla loro applicazione, nonché quello di confrontare le misure che ne risultano.

REFERENCES

Andrich D. (1978). A rating formulation for ordered response categories. *Psychometrika*, **43**, 561-573.

Bartholomew D.J. (1996). *The statistical approach to the social measurement*. Academic Press Inc., London.

Borzaga C. (2000). *Capitale umano e qualità del lavoro nei servizi sociali*. FIVOL, Roma.

Borzaga C., Musella M. (2003). *Produttività ed efficienza nelle organizzazioni nonprofit. Il ruolo dei lavoratori e delle relazioni di lavoro*, Edizioni 31, Trento.

Brentari E., Golia S. (2005). The Rasch model and the FIVOL-FEO survey. *Rapporti di Ricerca del Dipartimento di Metodi Quantitativi,* Brescia University, Working paper 257.

Carpita M., Manisera M. (2006). Un'analisi delle relazioni tra equità, motivazione e soddisfazione per il lavoro. In M. Carpita, L. D'Ambra, M. Vichi, G. Vittadini (eds.), *Valutare la qualità. I servizi di pubblica utilità alla persona*. Guerini editore, Milano.

Ferrari A., Annoni P., Salini S. (2005). A comparison between alternative models for environmental ordinal data: Nonlinear PCA vs Rasch Analysis. *Proceedings of the 20th International Workshop on Statistical Modelling*, Sydney, 173-177.

Gifi A. (1990). *Nonlinear multivariate analysis*. Wiley, Chichester.

Heiser W.J., Meulman J.J. (1994). Homogeneity analysis: exploring the distribution of variables and their nonlinear relationships. In Greenacre M., Blasius J. (eds.), *Correspondence analysis in the social sciences: recent developments and applications*. Academic Press, Harcourt Brace & Company, New York.

Lin S.P., Bendel R.B. (1985). Algorithm AS213: Generation of population correlation on matrices with specified eigenvalues. *Applied Statistics*, **34**, 193-198.

Linacre J.M. (2005). *A user's guide to WINSTEPS* [Computer program, version 3.57]. Chicago.

Linting M., Meulman J.J., Groenen P.J.F., Van der Kooij A.J. (2007). Nonlinear Principal Components Analysis: Introduction and application. *Psychological Methods.* In press.

Manisera M. (2005). Sulla scelta dello scaling level nell'Analisi delle Componenti Principali Nonlineare. In Amenta P., D'Ambra L., Squillante M., Ventre A. (eds.), *Metodi, modelli e tecnologie dell'informazione a supporto delle decisioni*. Franco Angeli, Milano.

Manisera M., Dusseldorp E.M.L.,Van der Kooij A.J. (2005). Component structure of job satisfaction based on Herzberg's theory. *Rapporti di Ricerca del Dipartimento di Metodi Quantitativi,* Brescia University, Working paper 253.

Masters G.N. (1982). A Rasch model for partial credit scoring. *Psychometrika*, **47**, 149-174.

Meulman J.J. (1982). *Homogeneity analysis of incomplete data*. DSWO Press, Leiden.

Meulman J.J., Van der Kooij A.J., Heiser W.J. (2004). Principal Component Analysis with nonlinear optimal scaling transformations for ordinal and nominal data. In Kaplan D. (ed.), *The Sage Handbook of Quantitative Methodology for the Social Sciences*. Sage, London.

Michailidis G., de Leeuw J. (1998). The Gifi system of descriptive multivariate analysis. *Statistical Science*, **13**, 307-336.

Ramsay J.O. (1988). Monotone Regression Splines in Action. *Statistical Science*, **3**, 425-461.

Rasch G. (1960). *Probabilistic models for some intelligence and attainment tests*. The Danish Institute of Educational Research, Copenhagen.

Smith Jr. E.V. (2002). Detecting and evaluating the impact of multidimensionality using item fit statistics and principal component analysis of residuals. *Journal of Applied Measurement*, **3**, 205-231.

SPSS (2003). CatPCA algorithm, Retrieved May 20, 2004. http://support.spss.com/.

van Rijckevorsel J., Bettonvil B., de Leeuw J. (1985). *Recovery and stability in nonlinear PCA*. Department of Data Theory, University of Leiden, RR-85-21.

Stevens J. (1992). *Applied multivariate statistics for the social sciences (2nd ed.)*. Lawrence Erlbaum Associates, Hillsdale.

Wright B.D., Masters G.N. (1982). *Rating Scale Analysis*. MESA Press, Chicago.

Appendix

The 13 items of the satisfaction section considered in the analysis

Item	
	How satisfied are you ...
1	with your formative training and professional growth at this organization?
2	with your decisional and operative independence?
3	with the recognition by the others of your work done?
4	with the variety and creativity of the work?
5	with the working environment (safety, comfort)?
6	with the benefit that your work produces for end users?
7	with your pay?
8	with the working hours schedule?
9	with the career promotions achieved up to this moment in this organization?
10	with the promotion prospects in this organization?
11	with the certainty of your job?
12	with the relation with your superiors?
13	with the relation with your paid co-workers?