Qual Quant (2010) 44:153–166
DOI 10.1007/s11135-008-9190-y

RESEARCH NOTE

Polychoric versus Pearson correlations in exploratory and confirmatory factor analysis of ordinal variables

Francisco Pablo Holgado –Tello · Salvador Chacón –Moscoso · Isabel Barbero –García · Enrique Vila –Abad

Published online: 7 September 2008

Abstract Given that the use of Likert scales is increasingly common in the field of social research it is necessary to determine which methodology is the most suitable for analysing the data obtained; although, given the categorization of these scales, the results should be treated as ordinal data it is often the case that they are analysed using techniques designed for cardinal measures. One of the most widely used techniques for studying the construct validity of data is factor analysis, whether exploratory or confirmatory, and this method uses correlation matrices (generally Pearson) to obtain factor solutions. In this context, and by means of simulation studies, we aim to illustrate the advantages of using polychoric rather than Pearson correlations, taking into account that the latter require quantitative variables measured in intervals, and that the relationship between these variables has to be monotonic. The results show that the solutions obtained using polychoric correlations provide a more accurate reproduction of the measurement model used to generate the data.

Keywords Construct validity · Polychoric correlations · Pearson correlation · Factor analysis

The present study forms part of the results obtained in research project SEJ2004-05360/EDUC funded by Spain's Ministerio de Educación y Ciencia.

F. P. Holgado–Tello (✉) · I. Barbero–García · E. Vila–Abad
Dpto. Metodología de las Ciencias del Comportamiento, Facultad de Psicología,
UNED, c/ Juan del Rosal, no 10, 28040 Madrid, Spain
e-mail: pfholgado@psi.uned.es

I. Barbero–García
e-mail: mbarbero@psi.uned.es

E. Vila–Abad
e-mail: evila@psi.uned.es

S. Chacón–Moscoso
Metodología de las Ciencias del Comportamiento, Facultad de Psicología,
Universidad de Sevilla,
c/ Camilo José Cela, 41018 Sevilla, Spain
e-mail: schacon@us.es

1 Introduction

Although Likert scales are widely used to obtain data, when studying the dimensionality of such data and gathering evidence about their construct validity, both exploratory factor analysis (EFA) and confirmatory factor analysis (CFA) may be conducted using Pearson correlation matrices. In this study, by means of simulation studies and taking into account the metric derived from this kind of ordinal scale, we aim to demonstrate the advantages of polychoric correlations over Pearson correlations when carrying out this kind of analysis.

There are several reasons why Pearson correlations can be considered a less suitable method for studying the degree of association between categorical variables. First, from a methodological point of view these variables would imply ordinal scales, whereas Pearson correlations assume interval measurement scales. Furthermore, the only information provided by this kind of scale is the number of subjects in each of the categories (cells) in a contingency table; if Pearson correlations are used in this case the relationship between measures would be artificially restricted due to the restrictions imposed by categorization (Guilley and Uhlig 1993), since all subjects situated in the interval that limits each of the categories would be considered as being included in the same category and, therefore, they would be assigned the same score with a resulting reduction in data variability. Given that Pearson correlations decrease with the homogeneity of the sample, these restrictions in assigning scores to subjects would lead to an underestimate of the degree of association between observed variables and, consequently, a decrease in the factor weightings obtained from the factorization of the correlation matrix (Saris et al. 1998; DiStefano 2002). Therefore, when factor analysis is used to test the construct validity of a measurement instrument, it is important to take into account the measurement scale that is being used (Maydeu and D'Zurilla 1995; Jöreskog 2001; Flora et al. 2003).

In attempting to solve this problem, Jöreskog and Sörbom (1996a) found that regardless of sample size and population correlation, polychoric correlations were the most consistent and robust estimator.

Dollan (1994) argued that at least five response categories are required to apply the normal factor analytic theory to Pearson correlations. However, an underestimate of both parameter estimates and standard error is observed even under the ideal circumstances in which skewness and kurtosis do not depart from their ideal values under normality.

It should also be noted that methods of estimation become especially relevant in confirmatory factor analysis (CFA) when investigating the relationship between ordinal variables using structural equation models. If the variables are ordinal the relationships between them should be analysed using polychoric correlations, where the variance–covariance matrix, used as a weighting element, becomes an essential element within the estimation process (Bollen 1989; DiStefano 2002).

Taking into account (as pointed out by Kampen and Swyngedouw 2000; Cliff and Kyats 2005; Göb et al. 2007) that the use of measurement instruments which require categorical responses from subjects is increasingly common in social research, and that this implies the use of ordinal scales, the present study aims to demonstrate, by means of simulation studies, that the use of polychoric correlations with this kind of variable provides a more accurate reproduction of the model used to generate the data than do Pearson correlations.

2 The basis of polychoric correlations

Briefly, let us suppose that Z_1 and Z_2 are two ordinal items with m_1 and m_2 categories. Their distribution in the sample is given by the contingency table

$$\begin{array}{llll} n_{11} & n_{12} & \cdots & n_{1m2} \\ n_{21} & n_{22} & \cdots & n_{2m2} \\ . & . & . & . \\ . & . & n_{ij} & . \\ . & . & . & . \\ n_{m11} & n_{m11} & \cdots & n_{m1m2} \end{array}$$

where n_{ij} is the number of cases in category i of item 1 and in category j of item 2. If we suppose that underlying these items are variables $Z^*{}_1$ and $Z^*{}_2$, which are normally distributed, it can be assumed that their combined distribution is a normal bivariate distribution with a correlation ρ. The polychoric correlation is the correlation ρ in the bivariate normal distribution N(0,0,1,1,ρ) (Eq. 1) of the latent variables $Z^*{}_1$ and $Z^*{}_2$. If $m_1 = m_2 = 2$ then the correlation is tetrachoric.

$$P[X = i, Y = j] = p_{ij} = \int_{a_{i-1}}^{a_i} \int_{b_{j-1}}^{b_j} \frac{1}{2\pi\sqrt{1-\rho^2}} \exp^{-\frac{1}{2(1-\rho^2)}(x^2-2\rho xy+y^2)} dxdy \quad (1)$$

and can be estimated by maximizing the function of maximum likelihood of the multinomial distribution:

$$\ln L = \sum_{i=1}^{m1} \sum_{j=1}^{m2} n_{ij} \log p_{ij} \quad (2)$$

The polychoric correlation is used when the variables are continuous and linearly related but, contrary to the requirements of Pearson correlations, are divided into a series of categories.

Although, in theory, it is necessary to test the assumption of bivariate normality before calculating the polychoric correlation, this correlation is fairly robust with respect to such a violation (Coenders et al. 1997). Furthermore, given the sensitivity of chi-squared tests, particularly in large samples, it is necessary to find alternative statistics for evaluating the assumption of normality. Jöreskog (2001) proposes using the root mean square error of approximation (RMSEA) as a fit index, as when its values are no greater than 0.1 parameter estimation is not significantly affected, even when the variables do not show bivariate normality.

3 Categorisation in factor analysis

One of the techniques able to provide evidence of the construct validity of a test or scale, and which has been used in the present study, is factor analysis. This technique reveals the internal structure of the covariance or correlation matrices.

As pointed out above, if Pearson correlations are used to analyse the validity of the scale and quantify the degree of association between ordinal variables, the values obtained will be lower as all subjects situated at different points of the interval will be assigned the same score. However, if the subjects were able to be situated along the latent continuum, without the category restrictions, the values obtained may be different because Pearson correlations reduce the magnitude of the coefficients obtained among observed variables due to the fact that the categorisation reduces variability; therefore, problems of estimation may arise (Guilley and Uhlig 1993). Consequently, the factor loadings obtained when factoring the correlation matrix will also be reduced as there is not only a random error but also a category error effect (Saris et al. 1998; DiStefano 2002). It is therefore important, when using factor analysis to test construct validity, to take into account the type of scale used for measuring the observable variables (Maydeu and D'Zurilla 1995; Jöreskog 2001; Flora et al. 2003).

It can be seen from the above that the metric of the data influences how they should be analysed. Thus, when using Likert-type items and investigating the relationship between them by means of structural equation models, the methods of estimation employed become particularly important (DiStefano 2002). The most popular among estimators based on normal distributions is the maximum likelihood (ML) method, as this yields consistent and asymptotically unbiased parameters (Bollen 1989). However, if the variables are ordinal the relationships between them should be analysed using polychoric correlations. In this procedure, the weighted least squares (WLS) method, a particular case of the generalised least squares (GLS) procedure, is recommended when sample size is large (Jöreskog and Sörbom 1996b; Flora and Curran 2004). In fact, a simulation study carried out by DiStefano (2002) found that WLS showed a small bias in estimating parameters and that this bias was reduced as sample size increased.

4 Procedure

First of all we generated, assuming a normal distribution N (0,1), the answers of three samples (each comprising 1,000 subjects) to 12 items following a theoretical model of three, four and five dimensions, respectively. Second, we categorized the answers on a five-point scale so that: (a) the distribution of the answers to all items, except one, was symmetric; and (b) skewness was introduced in all items. Asymmetry was alternatively positive and negative.

In this way, the limits for categorization were as follows:

(1) In the symmetric items:

$$\begin{array}{llll} \text{If} & & z^* \leq -1.5 & z \text{ is codified as } 1 \\ \text{If } -1.5 & < & z^* \leq -0.5 & z \text{ is codified as } 2 \\ \text{If } -0.5 & < & z^* \leq 0.5 & z \text{ is codified as } 3 \\ \text{If } 0.5 & < & z^* \leq 1.5 & z \text{ is codified as } 4 \\ \text{If } 1.5 & < & z^* & z \text{ is codified as } 5 \end{array}$$

(2) In the negative symmetric items:

$$\begin{array}{llll} \text{If} & & z^* \leq 1 & z \text{ is codified as } 1 \\ \text{If } 1 & < & z^* \leq 1.5 & z \text{ is codified as } 2 \\ \text{If } 1.5 & < & z^* \leq 2 & z \text{ is codified as } 3 \\ \text{If } 2 & < & z^* \leq 2.5 & z \text{ is codified as } 4 \\ \text{If } 2.5 & < & z^* & z \text{ is codified as } 5 \end{array}$$

(3) In the positive symmetric items:

$$\begin{array}{llll} \text{If} & & \text{z}^* \leq -2.5 & \text{z is codified as } 1 \\ \text{If } -2.5 & < & \text{z}^* \leq -2 & \text{z is codified as } 2 \\ \text{If } -2 & < & \text{z}^* \leq -1.5 & \text{z is codified as } 3 \\ \text{If } -1.5 & < & \text{z}^* \leq 1 & \text{z is codified as } 4 \\ \text{If } 1 & < & \text{z}^* & \text{z is codified as } 5 \end{array}$$

One hundred replications were carried out for each experimental condition.

Third, an exploratory factor analysis (EFA) was carried out, from the matrices of both Pearson and polychoric correlations, in order to analyse how the theoretical models were reproduced. Finally, a confirmatory factor analysis (CFA) was used to determine the goodness of fit of the models with both kinds of correlations.

In summary, six experimental conditions were generated as a result of the combination of two factors: the number of dimensions correlated (3, 4 and 5) and the number of items with asymmetry: only one asymmetric item or all asymmetric items (alternatively with positive and negative asymmetry). The number of replications in each of the experimental conditions was 100, the number of items 12 and the sample size in each replication was 1,000 subjects.

The resulting matrices Λ_y with the factor saturations for the generation of data following the theoretical model, as well as Ψ matrixes were:

$$\Lambda_y = \begin{pmatrix} .7 & 0 & 0 \\ .6 & 0 & 0 \\ .5 & 0 & 0 \\ .4 & 0 & 0 \\ 0 & .7 & 0 \\ 0 & .6 & 0 \\ 0 & .5 & 0 \\ 0 & .4 & 0 \\ 0 & 0 & .7 \\ 0 & 0 & .6 \\ 0 & 0 & .5 \\ .4 & .4 & .4 \end{pmatrix} \quad \Lambda_y = \begin{pmatrix} .7 & 0 & 0 & 0 \\ .6 & 0 & 0 & 0 \\ .5 & 0 & 0 & 0 \\ 0 & .7 & 0 & 0 \\ 0 & .6 & 0 & 0 \\ 0 & .5 & 0 & 0 \\ 0 & 0 & .7 & 0 \\ 0 & 0 & .6 & 0 \\ 0 & 0 & .5 & 0 \\ 0 & 0 & 0 & .7 \\ 0 & 0 & 0 & .6 \\ .5 & .5 & .5 & .5 \end{pmatrix} \quad \Lambda_y = \begin{pmatrix} .7 & 0 & 0 & 0 & 0 \\ .6 & 0 & 0 & 0 & 0 \\ .5 & .5 & .5 & .5 & .5 \\ 0 & .7 & 0 & 0 & 0 \\ 0 & .6 & 0 & 0 & 0 \\ 0 & .5 & 0 & 0 & 0 \\ 0 & 0 & .7 & 0 & 0 \\ 0 & 0 & .6 & 0 & 0 \\ 0 & 0 & 0 & .7 & 0 \\ 0 & 0 & 0 & .6 & 0 \\ 0 & 0 & 0 & 0 & .7 \\ 0 & 0 & 0 & 0 & .6 \end{pmatrix}$$

$$\Psi = \begin{pmatrix} 1 & & \\ .456 & 1 & \\ .673 & .309 & 1 \end{pmatrix} \quad \Psi = \begin{pmatrix} 1 & & & \\ .353 & 1 & & \\ .510 & .177 & 1 & \\ .628 & .223 & .321 & 1 \end{pmatrix} \quad \Psi = \begin{pmatrix} 1 & & & & \\ .367 & 1 & & & \\ .501 & .171 & 1 & & \\ .627 & .231 & .324 & 1 & \\ .662 & .254 & .319 & .405 & 1 \end{pmatrix}$$

It can be seen that in each of the matrices that include factor saturation, there is an item that saturates on all factors. This item is asymmetric in all cases and its asymmetry is positive in the model of three and five dimensions and negative in the model of four dimensions.

These experimental conditions were chosen in order to be able to compare the results with those obtained in previous studies carried out by our group (Holgado et al. 2006, 2007). However, we wish to consider a wide range of situations and aim to change the experimental conditions in future studies.

The software used for data generation and analysis was PRELIS 2.30 and LISREL 8.54 (Jöreskog and Sörbom 1996a,b), along with SPSS 14.0; the programming language was obtained through online free shareware: *R 2.0.0 "A Language and Environment"* (2004).

After generating the data we conducted an oblique exploratory factor analysis (EFA), with maximum likelihood (ML), and a confirmatory factor analysis (CFA), with weighted least squares (WLS) and maximum likelihood (ML), using polychoric and Pearson correlations, respectively (Flora et al. 2003; Flora and Curran 2004).

5 Results

5.1 Main results from the exploratory factor analysis

Table 1 shows both polychoric and Pearson correlation matrices obtained with data generated for the model of three dimensions and with one asymmetric item per dimension. This item would be: in the first dimension, number 4 with positive asymmetry; in the second dimension, number 8 with negative asymmetry; and number 12 that saturates on all dimensions and shows positive asymmetry. In Table 2 the same matrices are shown, but under the condition that all items were asymmetric; in this case, items with positive and negative asymmetry would be alternated from the first item that presented positive asymmetry.

As can be seen, Pearson correlations (values above the diagonal) are lower then polychoric correlations (values below the diagonal) when only one asymmetric item has been generated, as well as when all items are asymmetric. In the latter case, while polychoric correlation values do not differ very much, Pearson correlation values show a clear drop.

With respect to the exploratory factor analysis, when using the polychoric correlation matrix we obtained the same three-factor structure as in the theoretical model used for data generation, regardless of whether there is only one asymmetric item or all items are

Table 1 Matrix of polychoric (below diagonal) and Pearson (above) correlations for 3F (one item skewness)

	1	2	3	4	5	6	7	8	9	10	11	12
1	–	0.385	0.354	.150	0.183	0.172	0.156	0.125	0.306	0.296	0.279	0.312
2	0.422	–	0.320	0.131	0.186	0.168	0.160	0.128	0.286	0.287	0.257	0.300
3	0.388	0.351	–	0.134	0.166	0.164	0.134	0.118	0.270	0.243	0.256	0.287
4	0.331	0.304	0.301	–	0.078	0.069	0.068	0.037	0.109	0.101	0.091	0.153
5	0.204	0.207	0.185	0.181	–	0.443	0.402	0.328	0.143	0.140	0.118	0.272
6	0.190	0.186	0.180	0.155	0.485	–	0.390	0.297	0.145	0.128	0.115	0.272
7	0.174	0.177	0.148	0.152	0.441	0.426	–	0.285	0.121	0.121	0.098	0.252
8	0.158	0.156	0.147	0.108	0.408	0.369	0.354	–	0.096	0.085	0.079	0.158
9	0.339	0.317	0.298	0.241	0.161	0.162	0.137	0.121	–	0.507	0.482	0.351
10	0.327	0.318	0.268	0.224	0.157	0.142	0.134	0.106	0.557	–	0.456	0.337
11	0.308	0.285	0.281	0.204	0.131	0.124	0.109	0.098	0.530	0.500	–	0.311
12	0.484	0.462	0.446	0.363	0.420	0.386	0.387	0.331	0.541	0.521	0.481	–

Table 2 Matrix of polychoric (below diagonal) and Pearson (above) correlations for 3F, all item skewness)

	1	2	3	4	5	6	7	8	9	10	11	12
1	–	0.142	0.179	0.106	0.086	0.069	0.062	0.049	0.144	0.117	0.124	0.157
2	0.460	–	0.106	0.253	0.077	0.148	0.073	0.139	0.123	0.259	0.099	0.395
3	0.399	0.359	–	0.092	0.060	0.051	0.043	0.039	0.100	0.088	0.106	0.121
4	0.311	0.318	0.293	–	0.073	0.123	0.053	0.112	0.115	0.196	0.090	0.308
5	0.225	0.182	0.166	0.177	–	0.155	0.196	0. 142	0.061	0.113	0.066	0.329
6	0.184	0.186	0.148	0.157	0.475	–	0.134	0.303	0.061	0.113	0.066	0.329
7	0.186	0.203	0.153	0.137	0.435	0.428	–	0.114	0.031	0.046	0.020	0.120
8	0.144	0.174	0.113	0.140	0.415	0.372	0.349	–	0.044	0.082	0.039	0.261
9	0.331	0.304	0.258	0.273	0.162	0.138	0.123	0.090	–	0.192	0.293	0.196
10	0.341	0.323	0.274	0.250	0.154	0.150	0.116	0.109	0.560	–	0.159	0.436
11	0.305	0.273	0.275	0.235	0.118	0.155	0.063	0.105	0.533	0.507	–	0.156
12	0.503	0.483	0.399	0.383	0.429	0.409	0.351	0.332	0.533	0.528	0.487	–

Table 3 Lambda matrix for three factors, one item skewness (in brackets lambdas for all items skewness)

Item	Polychoric correlations			Pearson correlations		
	F1	*F2*	*F3*	*F1*	*F2*	*F3*
1	**0.660(0.693)**	−0.024(−0.020)	0.026(0.008)	0.277(0.211)	**0.428**(0.142)	**0.645**(0.289)
2	**0.602(0.663)**	0.003(0.000)	0.027(−0.011)	0.278(**0.495**)	**0.405**(0.268)	**0.589**(0.216)
3	**0.579(0.582)**	−0.008(−0.034)	0.017(0.004)	0.254(0.164)	**0.376**(0.108)	**0.549**(0.235)
4	**0.514(0.426)**	0.009(0.033)	−0.029(0.068)	0.117(**0.392**)	0.152(0.220)	0.240(0.194)
5	0.032(0.007)	**0.704**(**0.707**)	−0.012(0.006)	**0.683**(0.159)	0.213(**0.327**)	**0.317**(0.112)
6	0.013(−0.029)	**0.662**(**0.674**)	0.001(0.040)	**0.645**(**0.344**)	0.205(**0.562**)	0.295(0.076)
7	0.011(0.065)	**0.627**(**0.609**)	−0.003(−0.067)	**0.600**(0.126)	0.183(0.285)	0.272(0.067)
8	0.009(0.005)	**0.562**(**0.571**)	−0.011(−0.017)	**0.468**(0.276)	0.135(**0.498**)	0.211(0.043)
9	0.070(0.028)	−0.004(−0.008)	**0.723**(**0.759**)	0.229(0.266)	**0.731**(0.099)	**0.481**(**0.542**)
10	0.073(0.092)	−0.010(−0.010)	**0.685**(**0.670**)	0.215(**0.545**)	**0.693**(0.183)	**0.463**(**0.311**)
11	0.078(0.054)	−0.034(−0.028)	**0.651**(**0.669**)	0.186(0.216)	**0.658**(0.089)	**0.440**(**0.505**)
12	**0.380(0.346)**	**0.317**(**0.333**)	**0.373**(**0.407**)	**0.403**(**0.812**)	**0.489**(**0.529**)	**0.512**(0.295)

In bold, the loadings higher 0.3

Table 4 Phi matrix for three factors

Dimension	1	2	3
1	–	0.332(0.558)	0.477(0.420)
2	0.435(0.433)	–	0.676(0.136)
3	0.590(0.593)	0.271(0.264)	–

Polychoric (below diagonal) and Pearson (above). (In brackets phi for all items skewness)

Table 5 Matrix of polychoric (below diagonal) and Pearson (above) correlations for 4F (one item skewness)

	1	2	3	4	5	6	7	8	9	10	11	12
1	–	0.390	0.178	0.140	0.147	0.118	0.216	0.211	0.112	0.286	0.271	0.419
2	0.428	–	0.165	0.143	0.146	0.115	0.203	0.193	0.095	0.275	0.273	0.405
3	0.395	0.363	–	0.062	0.055	0.040	0.089	0.079	0.069	0.129	0.120	0.137
4	0.154	0.158	0.142	–	0.424	0.344	0.087	0.072	0.048	0.106	0.111	0.314
5	0.161	0.160	0.119	0.463	–	0.323	0.068	0.053	0.055	0.100	0.097	0.286
6	0.147	0.137	0.106	0.431	0.399	–	0.058	0.059	0.023	0.087	0.086	0.263
7	0.240	0.225	0.210	0.095	0.076	0.069	–	0.461	0.219	0.156	0.148	0.376
8	0.233	0.215	0.177	0.079	0.057	0.075	0.504	–	0.217	0.131	0.133	0.338
9	0.230	0.195	0.202	0.097	0.112	0.065	0.453	0.445	–	0.069	0.060	0.137
10	0.316	0.303	0.284	0.120	0.112	0.105	0.174	0.146	0.144	–	0.497	0.431
11	0.299	0.299	0.265	0.126	0.108	0.105	0.164	0.149	0.123	0.545	–	0.421
12	0.517	0.500	0.444	0.387	0.351	0.336	0.461	0.417	0.398	0.530	0.520	–

asymmetric (in brackets). However, when we consider the Pearson correlation matrix, the data are not consistent with the structure used to generate data (Table 3).

When we analyze the correlation matrix (ψ) between factors, we can see that the estimation is better when we have used polychoric correlations (Table 4).

Tables 5 and 6 show polychoric and Pearson correlation matrices obtained when using four dimensions for data generation, regardless of whether we introduced one asymmetric item per dimension or all asymmetric items. In the first case, the asymmetric items would be number 3 in the first dimension (with positive asymmetry), number 6 in the second dimension (with negative asymmetry), number 9 in the third dimension (with positive asymmetry) and,

Table 6 Matrix of polychoric (below diagonal) and Pearson (above) correlations for 4F (all item skewness)

	1	2	3	4	5	6	7	8	9	10	11	12
1	–	0.132	0.157	0.052	0.057	0.045	0.082	0.091	0.082	0.137	0.100	0.312
2	0.396	–	0.109	0.067	0.115	0.054	0.087	0.168	0.073	0.116	0.224	0.255
3	0.395	0.364	–	0.037	0.050	0.053	0.055	0.063	0.070	0.097	0.087	0.243
4	0.153	0.169	0.136	–	0.151	0.185	0.035	0.035	0.022	0.037	0.047	0.212
5	0.136	0.152	0.128	0.471	–	0.123	0.024	0.048	0.030	0.052	0.084	0.192
6	0.161	0.139	0.146	0.427	0.396	–	0.025	0.012	0.001	0.030	0.038	0.148
7	0.247	0.202	0.176	0.099	0.059	0.101	–	0.165	0.224	0.082	0.066	0.260
8	0.243	0.213	0.185	0.093	0.062	0.044	501.	–	0.130	0.071	0.115	0.230
9	0.246	0.194	0.200	0.094	0.089	0.036	0.466	0.445	–	0.042	0.053	0.210
10	0.341	0.287	0.274	0.091	0.126	0.067	0.208	0.159	0.152	–	0.180	0.335
11	0.272	0.289	0.271	0.124	0.108	0.100	0.148	0.147	0.130	0.525	–	0.261
12	0.567	0.532	0.506	0.414	0.364	0.339	0.473	0.463	0.436	0.563	0.538	–

finally, number 12, which saturates on all factors and has negative asymmetry. As can be seen, the trend is the same as in the three-dimension model, that is, Pearson correlations are lower than polychoric correlations.

We then analysed the factor structures obtained from the previous matrices according to the criterion of considering only loadlugs higher than 0.30 and with factorization based on the polychoric correlation matrix. It can be seen that only item number 12, which shares factor saturations on the four factors, does not reach (in factor 2) the prefixed value, regardless of whether there is only one asymmetric item or all items are asymmetric; *the factor value shows good consistency with the model*. In contrast, when factorization is done using the Pearson correlation matrix the results are worse (Table 7).

The same trend is observed in the matrix that includes correlations (ψ) between factors, and the estimation is better when using polychoric correlations (Table 8).

The results obtained for the model of five dimensions show the same trend. The correlations obtained from the polychoric correlation matrix are higher than those from the Pearson correlation matrix, regardless of whether there is only one asymmetric item per dimension or all items are asymmetric. In the first case the asymmetric items are number 3, which saturates on all dimensions and has positive asymmetry, number 6, which saturates on the second dimension and has negative asymmetry, number 8 in the third dimension and with positive asymmetry, number 9 in the fourth dimension and with negative asymmetry and, finally, number 12, which saturates on the fifth dimension and has positive asymmetry (Tables 9, 10).

When we analyse the factor values obtained from the polychoric correlation matrix, the loadings obtained for item 3, which shares saturation on all factors, does not reach in the first three factors the value of 0.30 set as a criterion. In all the other items the theoretical model is reproduced regardless of whether there is only one asymmetric item or all items are asymmetric. When the factorization is done using the Pearson correlation matrix, the fit to the theoretical model is worse (Table 11).

The same trend is observed in the matrix that includes correlations (ψ) between factors, and the estimation is better when using polychoric correlations (Table 12).

If we consider only the data obtained from polychoric correlations it can be seen that as the number of factors increases the fit between the data and the theoretical model worsens. For three factors we obtained a lambda matrix with the same structure as the original, but with four and five factors there were some differences (especially for five factors).

Table 7 Lambda matrix for 4F, 1 item skewness (in paragraph lambdas for all items skewness)

Item	Polychoric correlations				Pearson correlations			
	$F1$	$F2$	$F3$	$F4$	$F1$	$F2$	$F3$	$F4$
1	**0.708(0.691)**	−0.018(−0.038)	−0.007(0.008)	−0.023(−0.008)	0.275(**0.365**)	**0.360**(0.222)	**0.428**(0.176)	**0.652(0.185)**
2	**0.625(0.549)**	−0.013(−0.002)	−0.007(0.014)	−0.027(0.061)	0.274(**0.309**)	**0.336**(0.224)	**0.421**(0.206)	**0.604(0.595)**
3	**0.564(0.622)**	−0.013(−0.017)	0.001(−0.032)	0.019(−0.004)	0.099(0.286)	0.137(0.166)	0.176(0.147)	0.261(0.156)
4	−0.004(0.002)	**0.711(0.717)**	−0.006(0.023)	−0.003(−0.005)	**0.667**(0.192)	0.168(0.109)	0.204(**0.482**)	0.251(0.071)
5	0.032(−0.002)	**0.653(0.651)**	−0.020(−0.015)	−0.028(0.029)	**0.621**(0.195)	0.140(0.098)	0.184(**0.330**)	0.253(0.168)
6	0.006(0.089)	**0.615(0.584)**	−0.020(−0.037)	−0.004(−0.049)	**0.518**(0.138)	0.129(0.058)	0.167(**0.364**)	0.208(0.063)
7	0.008(−0.013)	−0.024(0.005)	**0.718(0.717)**	0.016(0.032)	0.175(0.256)	**0.704(0.502)**	0.275(0.122)	**0.350**(0.117)
8	0.003(0.046)	−0.032(−0.026)	**0.709(0.682)**	−0.011(−0.014)	0.151(0.247)	**0.653(0.337)**	0.240(0.120)	**0.333**(0.264)
9	0.054(0.086)	0.009(−0.015)	**0.616(0.616)**	−0.044(−0.040)	0.084(0.206)	**0.307(0.427)**	0.107(0.088)	0.168(0.099)
10	0.083(0.108)	−0.031(−0.045)	−0.016(0.008)	**0.698(0.679)**	0.220(**0.408**)	0.265(0.177)	**0.706**(0.147)	**0.462**(0.192)
11	0.044(0.073)	−0.023(−0.001)	−0.019(−0.028)	**0.725(0.680)**	0.219(**0.324**)	0.258(0.166)	**0.694**(0.155)	**0.446(0.369)**
12	**0.318(0.434)**	0.281(0.276)	**0.302(0.304)**	**0.353(0.341)**	**0.563(0.865)**	**0.578(0.604)**	**0.657(0.562)**	**0.668(0.360)**

In bold, the loadings higher 0.3

Table 8 Phi matrix for four factors

Dimension	1	2	3	4
1	–	0.614(0.511)	0.554(0.480)	0.452(0.625)
2	0.357(0.729)	–	0.334(0.516)	0.309(0.347)
3	0.465(0.561)	0.167(0.664)	–	0.258(0.464)
4	0.552(0.472)	0.209(0.566)	0.279(0.450)	–

Polychoric (below diagonal) and Pearson (above). (In paragraph phi for all items skewness)

Table 9 Matrix of polychoric (below diagonal) and Pearson (above) correlations for 5F (1 item skewness)

	1	2	3	4	5	6	7	8	9	10	11	12
1	–	0.370	0.402	0.152	0.131	0.126	0.204	0.179	0.174	0.273	0.307	0.256
2	0.407	–	0.374	0.153	0.143	0.124	0.207	0.174	0.171	0.251	0.283	0.248
3	0.537	0.502	–	0.259	0.244	0.195	0.322	0.236	0.311	0.372	0.402	0.307
4	0.169	0.167	0.355	–	0.418	0.341	0.081	0.066	0.075	0.108	0.121	0.114
5	0.147	0.157	0.334	0.459	–	0.336	0.071	0.052	0.064	0.104	0.115	0.103
6	0.154	0.148	0.326	0.425	0.414	–	0.089	0.055	0.056	0.089	0.110	0.099
7	0.228	0.229	0.438	0.092	0.080	0.105	–	0.392	0.107	0.151	0.157	0.138
8	0.219	0.213	0.399	0.083	0.065	0.068	0.491	–	0.079	0.122	0.139	0.117
9	0.307	0.311	0.548	0.139	0.124	0.118	0.191	0.172	–	0.318	0.136	0.096
10	0.302	0.277	0.499	0.119	0.116	0.107	0.167	0.147	0.564	–	0.199	0.174
11	0.339	0.313	0.536	0.135	0.127	0.132	0.176	0.169	0.243	0.220	–	0.452
12	0.316	0.305	0.543	0.133	0.123	0.121	0.171	0.154	0.220	0.214	0.564	–

Table 10 Matrix of polychoric (below diagonal) and Pearson (above) correlations for 5F, all items skewness)

	1	2	3	4	5	6	7	8	9	10	11	12
1	–	0.134	0.284	0.054	0.067	0.024	0.117	0.088	0.142	0.129	0.112	0.118
2	0.405	–	0.278	0.055	0.123	0.032	0.099	0.167	0.131	0.121	0.259	0.117
3	0.552	0.506	–	0.157	0.197	0.090	0.241	0.233	0.323	0.283	0.296	0.276
4	0.145	0.140	0.337	–	0.149	0.177	0.031	0.017	0.055	0.033	0.042	0.027
5	0.164	0.156	0.335	0.461	–	0.117	0.032	0.051	0.057	0.063	0.101	0.050
6	0.087	0.099	0.234	0.410	0.380	–	0.031	0.018	0.001	0.006	0.034	0.025
7	0.275	0.250	453.	0.086	0.071	0.077	–	0.168	0.082	0.091	0.063	0.058
8	0.235	0.214	0.400	0.047	0.068	0.052	0.503	–	0.078	0.070	0.128	0.064
9	0.336	0.327	0.554	0.145	0.129	0.013	0.216	0.174	–	0.317	0.104	0.083
10	0.307	0.313	0.520	0.104	0.132	0.036	0.228	0.167	0.566	–	0.090	0.083
11	0.325	0.324	0.538	0.099	0.134	0.096	0.157	0.167	0.241	0.220	–	0.187
12	0.317	0.288	0.502	0.095	0.116	0.088	0.126	0.150	0.199	0.208	0.564	–

5.2 Main results from the confirmatory factor analysis

Through the exploratory factor analysis we were able to show that when factorization is done using the polychoric correlation matrix, the results, within all experimental conditions, are more consistent with the theoretical model than when the Pearson correlation matrix is used. However, in order to determine the goodness of fit between the hypothesized model and the results obtained, a confirmatory factor analysis was also carried out.

Table 13 shows the results obtained under the different conditions.

Table 11 Lambda matrix for 5F, 1 item skewness (in paragraph lambdas for all items skewness)

Item	With polychoric correlations					With Pearson correlations				
	*F*1	*F*2	*F*3	*F*4	*F*5	*F*1	*F*2	*F*3	*F*4	*F*5
1	**0.716**(**0.637**)	−0.019(−0.015)	−0.017(0.002)	−0.025(0.014)	0.000(0.028)	0.256(0.277)	**0.400**(**0.329**)	**0.307**(**0.307**)	**0.417**(0.181)	**0.636**(0.191)
2	**0.535**(**0.578**)	0.006(−0.012)	0.012(0.012)	0.055(0.014)	0.035(0.031)	0.259(0.256)	**0.371**(**0.332**)	**0.304**(**0.323**)	**0.393**(0.206)	**0.580**(**0.551**)
3	0.198(0.206)	0.215(0.210)	0.245(0.237)	**0.303**(0.026)	**0.322**(**0.320**)	**0.425**(**0.622**)	**0.515**(**0.774**)	**0.452**(**0.689**)	**0.635**(**0.488**)	**0.650**(**0.451**)
4	0.013(−0.008)	**0.690**(**0.716**)	−0.012(−0.023)	−0.009(**0.320**)	−0.016(−0.042)	**0.651**(0.115)	0.175(0.148)	0.131(0.118)	0.207(**0.463**)	0.270(0.068)
5	−0.014(0.039)	**0.683**(**0.642**)	−0.024(−0.039)	−0.002(0.047)	−0.007(−0.005)	**0.643**(0.143)	0.164(0.212)	0.115(0.162)	0.192(**0.350**)	0.244(0.197)
6	0.000(−0.028)	**0.624**(**0.604**)	0.006(0.035)	−0.018(0.027)	−0.002(0.029)	**0.519**(0.029)	0.153(0.089)	0.127(0.078)	0.157(**0.345**)	0.215(0.049)
7	−0.018(0.030)	−0.008(−0.015)	**0.750**(**0.760**)	−0.004(−0.104)	−0.010(−0.058)	0.162(0.191)	0.222(0.230)	**0.818**(**0.424**)	0.277(0.132)	**0.33**(0.128)
8	0.025(0.014)	−0.032(−0.035)	**0.675**(**0.677**)	−0.021(0.012)	−0.015(0.038)	0.118(0.170)	0.189(0.235)	**0.484**(**0.385**)	0.208(0.115)	0.289(0.272)
9	−0.017(−0.001)	−0.016(−0.013)	−0.006(−0.027)	**0.819**(−0.026)	−0.013(−0.016)	0.134(**0.613**)	0.184(**0.319**)	0.164(0.284)	**0.572**(0.169)	**0.300**(0.180)
10	0.040(0.034)	−0.022(−0.014)	−0.029(0.011)	**0.705**(**0.828**)	−0.011(0.004)	0.196(**0.505**)	0.270(.287)	0.230(0.268)	**0.569**(0.144)	**0.447**(0.164)
11	0.080(0.027)	−0.018(−0.025)	−0.016(−0.015)	−0.009(**0.676**)	**0.682**(**0.746**)	0.223(0.212)	**0.863**(**0.413**)	0.248(0.258)	**0.347**(0.186)	**0.473**(**0.483**)
12	−0.031(0.009)	−0.022(−0.031)	−0.021(−0.037)	−0.018(0.137)	**0.829**(**0.227**)	0.197(0.183)	**0.533**(**0.405**)	0.211(0.211)	0.270(0.137)	**0.410**(0.227)

In bold, the loadings higher 0.3

Table 12 Phi matrix for five factors

Dimension	1	2	3	4	5
1	–	0.311(0.644)	0.246(0.592)	0.359(,371)	0.451(0.361)
2	0.398(0.355)	–	0.340(0.718)	0.462(0.513)	0.624(0.578)
3	0.512(0.540)	0.233(0.200)	–	0.409(0.436)	0.501(0.479)
4	0.615(0.643)	0.292(0.227)	0.367(0.384)	–	0.693(0.308)
5	0.640(0.635)	0.304(0.267)	0.360(0.327)	0.425(0.396)	–

Polychoric (below diagonal) and Pearson (above).(In brackets phi for all items skewness)

Table 13 Fits indexes of CFA (in brackets fits indexes for CFA all items skewness)

Matrix used	Chi square	*P*	*df*	GFI	AGFI	RMSEA
	Three factors					
Pearson	86.35(1165.8)	0.001(0.000)	49	1.00(0.98)	1.00(0.97)	0.009(0.05)
Polychoric	51.81(42.01)	0.360(0.75)	49	1.00(1.00)	1.00(1.00)	0.002(0.000)
	Four factors					
Pearson	113.73(502.3)	0.000(0.000)	45	1.00(0.99)	1.00(0.98)	0.013(0.033)
Polychoric	38.18(30.43)	0.75(0.95)	45	1.00(1.00)	1.00(1.00)	0.000(0.000)
	Five factors					
Pearson	115.62(389.5)	0.000(0.000)	40	1.00(0.99)	1.00(0.99)	0.014(0.030)
Polychoric	29.92(32.38)	0.88(0.80)	40	1.00(1.00)	1.00(1.00)	0.000(0.000)

The results of the CFA show that when the polychoric correlation matrix is used, and regardless of the experimental condition, we must accept the hypothesis that the matrix estimated by the model examined is the same as the empirical one.

When considering the three- or four-dimension models the probability of accepting the model is higher when all items are asymmetric. For three dimensions the probabilities are 0.36 (one asymmetric item) as opposed to 0.75 (all asymmetric items). For the five-dimension model the model is accepted with a higher probability when there is only one asymmetric item (0.88 as opposed to 0.80). When using the Pearson correlation matrix to test with a CFA the theoretical models, the results show that these have to be rejected in all cases, regardless of the number of dimensions and asymmetric items.

Another interesting result is observed in relation to the value of chi-square (χ^2): when using the Pearson correlation matrix we found that as the number of factors increases the value of χ^2 also increases, whereas it decreases when the polychoric correlation matrix is used. In the first case the value increases from 86.35 (three dimensions) to 115.52 (five dimensions), while in the second case it decreases from 51.81 (three dimensions) to 29.92 (five dimensions).

In order to test the goodness of fit we included, in addition to χ^2, the most common global indexes: GFI, AGFI and RMSEA. Briefly, values higher than 0.9 are considered satisfactory for the first two indexes, and values lower than 0.08 for the third index (Bollen 1989).

With these criteria it can be seen that the RMSEA index has better results when using the polychoric correlation matrix. In fact, regardless of the number of dimensions and items with asymmetry, we obtained values near to 0 in all cases. In contrast, although the results with the Pearson correlation matrix are within acceptable limits, they are slightly higher when using the polychoric correlations. It can also be seen that as the number of dimensions increases, the values of RMSEA also increase: from 0.009 (three dimensions) to 0.014 (five dimensions).

As regards GFI and AGFI, when using polychoric correlations we obtained values of 1 in all cases, regardless of the number of dimensions and items with asymmetry. When the Pearson correlation matrix was used we also obtained values of 1, but only if we included one asymmetric item; when all items present asymmetry, values range from 0.97 (three dimensions) to 0.99 (five dimensions). At all events, the resulting values are good enough, regardless of the number of dimensions and items with asymmetry.

6 Discussion and conclusions

The present study aimed to show the importance of using the most suitable method of analysis given the level of measurement of the variables analysed. Although it is common to see data obtained by Likert scales and analysed through interval-based measures, methodologists need to ensure that any inferences made from the results obtained are as accurate as possible.

Furthermore, it should be noted that measurement in the social sciences often implies important degrees of error, both random and systematic, which may bias the estimates of relationships between the variables measured and, therefore, produce bias in the substantive conclusions. In this regard, one of the key problems relates to the consequences that use of the different types of correlations have on substantive and methodological research, as they may increase error and contribute to the misinterpretation of evidence of construct validity, the cornerstone of basic and applied research (Cronbach 1984; Messick 1994; APA, AERA, NCME 1999; Shadish et al. 2002).

Here we have shown that when construct validity is analysed according to ordinal data obtained from Likert scales, the factor results show a better fit to the theoretical model when the factorization is carried out using the polychoric rather than the Pearson correlation matrix.

This simulation study has sought to show, in a clear and precise way, the advantages of using one procedure rather than the other when the aim is to analyse construct validity from a factor point of view. Therefore, we have not placed undue emphasis on the power and effectiveness of the estimations according to the procedures used, as this aspect will be addressed in future studies. Our focus here has been to identify the type of correlation that yields a factor solution more in keeping with the original measurement model, as we believe this to be of great importance in terms of drawing correct substantive conclusions.

Our results can be summarised as follows:

1. When conducting an EFA: (a) regardless of the number of dimensions and items with skewness, Pearson correlations are lower than polychoric correlations. The results are more significant when all items are asymmetric. (b) The model obtained is more consistent with the original measurement model when we factorize using the polychoric correlation matrix. This result does not depend on the number of dimensions and asymmetric items.
2. When conducting a CFA: (a) goodness of fit measured by chi-squared always shows probabilities lower than the nominal level, that is, the model is accepted regardless of the number of dimensions and asymmetric items; and (b) other global indexes of goodness of fit, such as GFI, AGFI and RMSEA, are better when we use the polychoric correlation matrix, but values are generally good enough in both cases (Pearson and polychoric).

At all events, the factor results obtained (from both an exploratory and confirmatory point of view) when we use polychoric correlations reproduce better the measurement model present in the data, regardless of the number of factors.

Obviously, there are many issues which are not analysed by this simulation study (number of items, reliability, cross-loading, sample size, higher order factor, estimation method used, number of categories, etc.), and improvements are required in this regard. However, these aspects can now be examined in future research, the present study having provided an introduction to the problem using simulated data.

References

APA, AERA, NCME: Standards for Educational and Psychological Testing. Washington, Author (1999)

Bollen, K.A.: Structural Equations with Latent Variables. Wiley, New York (1989)

Coenders, G., Saris, W., Satorra, A.: Alternative approaches to structural equation modeling of ordinal data: a Monte Carlo study. Struct. Equ. Model. **4**(4), 261–282 (1997)

Cliff, N., Kyats, J.A.: Ordinal measurement in the behavioural sciences. Lawrence Erlbaum, New Jersey (2005)

Cronbach, L.: Essentials of Psychological Testing. Harper and Row, New York (1984)

DiStefano, C.: The impact of categorization with confirmatory factor analysis. Struct. Equ. Model. **9**(3), 327–346 (2002)

Dollan, C.V.: Factor analysis of variables with 2,3,4 and 7 response categories: a comparison of categorical variables estimators using stimulated data. Br. J. Math. Stat. Psychol. **47**, 309–326 (1994)

Flora, D.B., Curran, P.J.: An empirical evaluation of alternative methods of estimation for confirmatory factor analysis with ordinal data. Psychol Meth **9**, 466–491 (2004)

Flora, B.F., Finkel, E.J., Foshee, V.A.: Higher order factor structure of a self-control test: evidence from confirmatory factor analysis with polychoric correlations. Educ. Psychol. Meas. **63**(1), 112–127 (2003)

Göb, R., McCollin, C.H., Ramalhoto, M.F.: Ordinal methodology in the analysis of Likert scales Quality & Quantity (M.F.)

Guilley, W., Uhlig, G.: Factor analysis and ordinal data. Education **114**(2), 258–264 (1993)

Holgado, F.P., Chacón, S., Barbero, I., Sanduvete, S.: Training satisfaction rating scale: development of a measurement model using polychoric correlations. Eur. J. Psychol. Assess. **23**(4), 268–279 (2006)

Holgado, F.P., Carrasco, M.A., del Barrio, M.V., Chacón, S.: Factor analysis of the big five questionnaire using polychoric correlations in children. Qual Quant Int. J. Meth. (2007)

Jöreskog, K.G.: Analysis of ordinal variables 2. Cross-Sectional Data. Text of the workshop “Structural equation modelling with LISREL 8.51”. Friedrich-Schiller-Universität, Jena (2001)

Jöreskog, K.G., Sörbom, D.: LISREL 8: User’s reference guide. Scientific Software International, Chicago (1996a)

Jöreskog, K.G., Sörbom, D.: PRELIS 2: User’s reference guide. Scientific Software International, Chicago (1996b)

Kampen, J., Swyngedouw, M.: The ordinal controversy revisited. Qual. Quant. **34**, 87–102 (2000)

Maydeu, A., D’Zurilla, T.J.: A factor analysis of the social problem-solving inventory using polychoric correlations. Eur. J. Psychol. Assess. **11**(2), 98–107 (1995)

Messick, S: Foundations of validity meaning and consequences in psychological assessment. Eur. J. Psychol. Assess. **10**(1), 1–9 (1994)

Saris, W., Van Wijk, T., Scherpenzeel, A.: Validity and reliability of subjective social indicators. Soc. Indic. Res. **45**, 173–199 (1998)

Shadish, W.R., Cook, T., Campbell, D.: Experimental and quasi-experimental design for generalized causal inference. Houghton-Mifflin, Boston (2002)