

FUNDAMENTALS OF APPLICATION FACTOR ANALYSIS IN EDUCATION AND REHABILITATION

Senad Mehmedinović¹

Review paper

DOI: 10.21554/hrr.041708

Faculty of Education and Rehabilitation Sciences, University of Tuzla

Received: 18.01.2016

Accepted: 10.03.2017

ABSTRACT

Factor analysis is one of multivariate data processing methods, which studies the causal relationships of phenomena, that is, the cause of integration. In the introductory part of the paper, the basic definitions and interpretations regarding the factor analysis and the terms of multivariate methods, and some examples have been given in defining the manifest and latent, as explorative and confirmative examples. The justification for the application of factor analysis is elaborated in the main part of the paper with reference to the various authors who have dealt with this issue. Also, the paper presents the procedures of factor analysis, and presents tables and graphs showing the results necessary for interpretation. Given that for special education and rehabilitation a biopsychosocial approach is fundamental, factor analysis can be a powerful tool when studying interconnections of different phenomena. Its proper application by educators-rehabilitators, who act to this problem, may help in understanding the causes of connections of phenomena, and as such it helps in the development of a treatment for the prevention, education and rehabilitation of persons with disabilities.

Keywords: *Factor analysis, multivariate methods, special education-rehabilitation.*

INTRODUCTION

Factor analysis is one of multivariate data processing methods aimed at studying the interconnection of phenomenon and causes of connections between phenomenon. Multivariate method is often cited in literature as the multivariate analysis, which is correct, but errors occur when using the term multivariate method. Hasanbegović (2015) pointed at this, who states that in some books the term "multivariate method" is used terminologically wrong, and pointed out that the concept of variation or "variants of" is more proper in those situations, while the term "variants" which reflects options, is used incorrectly. Factor analysis was created as a "product" of thoughts of psychologists on intelligence and the greatest merit is attributed to Charles

Edward Spearman (1863-1945). He pointed in his work from 1904, while studying the correlation between the results of different tests of intelligence, to the possibility of expressing a simple model, that is that all common features of intelligence according to Spearman can be reduced to one general "g" factor and one specific "s" factor (Petrovic, 2013). Factor analysis aims to reeducate a larger number of manifest variables on latent dimensions called factors. Manifest variables (something that is obvious) also represent the results of characteristics that vary, which are obtained based on the application of measuring instruments, subtests or particles within subtests depending on what is the subject and objective of the research.

¹Correspondence to:

Senad Mehmedinović, Faculty of Education and Rehabilitation Sciences, University of Tuzla

Ul. Univerzitetska 1, 75000 Tuzla, B&H

Phone: 00387 62 901 229

E-mail: senad.mehmedinovic@gmail.com

For example, motor skills, which are evaluated by the Gross Motor Function Measure Test (GMFM), are a manifest variable. Or let's say, subtest walking, running and sitting of the Gross Motor Function Measure Test is a manifest variable. Also, individual items within subtest walking, running and sitting can be manifest variables. All what measures a given phenomenon that we observe can be called a manifest variable. Latent (hidden) variables are created as a linear combination of manifest variables and they are called factors. For example Hasanbegović, Mehmedinović and Mahmutović (2012) pointed out three factors in a paper called "Latent structure of motor abilities and skills of deaf children", the first factor being "The Factor of physical abilities and mobility" which is the most important in determining motor skills in the population of deaf children. Direct scientific contribution of factor analysis in this case is reflected in the fact that with the separated factor the time is "shorten" and it contributes to the proposal of measures and treatment for the population of deaf children. In this example, the exploratory factor analysis is applied when the number and structure of factors are not known in advance. In addition to exploratory factor analysis, factor analysis can also be confirmatory.

Mejovšek (2003) states that the confirmatory factor analysis is applied in cases when checking the hypothesis on the number and structure of factors for which already exist some empirical evidence.

For example, in order to evaluate the questionnaire "Psychological Inventory of Criminal Thinking Styles", Dolezal (2007) used confirmatory factor analysis, and he found that it has good measurement properties. However, the structure of the obtained factors are different from Walters structure (with whom it was compared), where Dolezal (2007) states that it is necessary to make certain changes to the instrument and to conduct additional research to better explain the settings of Walters theory of criminal lifestyle. From the above examples it can be concluded, and as Fulgosi (1988) also states in the strategies of using factor analysis, that confirmatory factor analysis appears as an objective test for a particular structural model or theory. Also, this is supported by the claim that confirmatory factor analysis is a powerful instrument for scientific validation of structural theory, and that it is a strong barrier to speculative models, theories and concepts that have no objective empirical confirmation (Fulgosi and Markovic, 1976 see Fulgosi 1979).

When to apply and what are the conditions for the application of factor analysis

When checking the set goals or approval (rejection) of research hypothesis, a question arises about the justification of the use of certain analyzes. To reasonably apply a certain analysis, or statistical test, dependence or independence of the sample, the types of applied variables, assessment of normal distribution of data, and determining a measure of symmetry and kurtosis and modalities of the curve by which the researcher defines the method of parametric or non-parametric statistics are all taken into account. All this also applies to the justification of the use of factor analysis with special emphasis on the variables used in the process. For example, if we are determining the motivation of young people with disabilities to continue their education, while not including variables related to success in school, self-respect, the support of parents, the community, etc., then factor analysis will not give the desired result or it will give a superficial result (as factor analysis is based on correlation). Here we are talking about the logic of the application of factor analysis, as well as experience of researchers what they would like to receive from the applied model. According to Fulgosi (1988) in the factor analysis we start from the data from the interconnection between the observed variables when these variables are expressed in points of deviation, ie., when all results and data are expressed as deviations from the arithmetic means of corresponding variables.

For each variable there must be a known standard deviation of results, and thus the variance of results that is equal to the square of the standard deviation is known (Fulgosi, 1988).

When it comes to the size of the sample for the application of factor analysis, the authors give different recommendations, but certainly all of them agree that the sample must belong to the group of large samples. Hasanbegović (2015) states that the ratio of the sample versus the applied variables must be at least 3:1. Suzić (2007) states that the number of data must be at least five times greater than the number of variables, while Nunnally (1978) recommended a ratio of 10:1. Fulgosi (1988) states that the deficiency in the application of factor analysis is an unstable correlation coefficient, and that it could be expected to be more stable when the number of participants should exceed 100 and that it must be at least five times greater than the number of variables. Tabachnick & Fidell (2007) and Pallant (2011) claim that five observations per item are sufficient in most cases.

Procedures in factor analysis

Factor analysis begins with intercorrelations of manifest variables, with the goal of gaining insight into the matrix of obtained coefficients ranged from -1 to +1. According to Tabachnik and Fidell (2007) one of the conditions that manifest variables are subjected to factor analysis is that their interconnectedness must be greater than 0.30, otherwise the data is not suitable for factor analysis. After insight into the intercorrelations of manifest variables, the results of the two tests justifiability of using factor analysis are calculated: Kaiser-Meyer-Olkin's (KMO) and Bartlett's test of sphericity, which is a measure of representativeness of the sample set of manifest variables, resulting in a ra-

tio of real and squared correlation and squared partial correlation. For the application of factor analysis to be justified, Bartlett's test of sphericity must be significant (≤ 0.05), and it is recommended that the value of the KMO test should be greater than 0.6 (Tabachnik & Fidell, 2007). In summary, as Hasanbegović (2015) stated, important information for the analysis in the correlation matrix are: the coefficients, the level of significance, the determinant (which should be greater than zero), KMO and Bartlett's test. After obtaining the coefficients, that is where the correlation of manifest variables is greater than 0.30, a KMO greater than 0.6, which is statistically significant, concludes that there is a justification of observed set of manifest variables to undergo factor analysis (Example results Table 1).

Table 1. KMO i Bartlett's test

Kaiser-Meyer-Olkin Measure of Sampling Adequacy.		.72
Bartlett's Test of	Approx. Chi-Square	216.94
Sphericity	Df	21
	Sig.	.000

Determinant= 0,33

After obtaining the results from the correlation matrix, we approach to the selection of the appropriate model, which in a statistical package (SPSS) has seven models: model of main components, model of unweighted least squares, model of least generalization, model of maximum likelihood, the model of the main axis of factors, model of alpha factorization and model of interpretation of factors. The most commonly used are two basic models: the model of common factor and component model. The model common factors analyzes the common variance of manifest variables that are in the factor analysis is called communality. The component model is mainly used in the factor analysis, which is a model of main components proposed by Harold Hotelling (1933). With the main

components method of Harold Hotelling we get a number of major components equal to the number of initial variables. The main components are calculated in succession at first the first one, then the second one, etc., where the first main component is counted on the full matrix of intercorrelation of variables and it describes the greatest amount of variance of the variable (Mejovšek, 2003). Having defined the model, the number of components that need to be kept are resolved using criteria. There are two key criteria and these are: Guttman-Kaiser criterion and scree test proposed by Cattell (1966), and it is called Cattell Scree Test. The Kaiser-Guttman criterion implies to take the factors or components with a characteristic value above one (Example Table 2).

Table 2. Total variance explained

Component	Initial Eigenvalues		
	Total	% of Variance	Cumulative %
1	2.53	36.16	36.16
2	1.06	15.21	51.38
3	.97	13.89	65.27
4	.76	10.94	76.22
5	.65	9.31	85.54
6	.55	7.96	93.50
7	.45	6.49	100.00

According to Cattell Scree test, as Pallant (2011) states, the characteristic values of all factors (SPSS works automatically) are drawn and recommended that all factors above the critical point should be kept, ie. the saddle of this diagram (Example Chart 1).

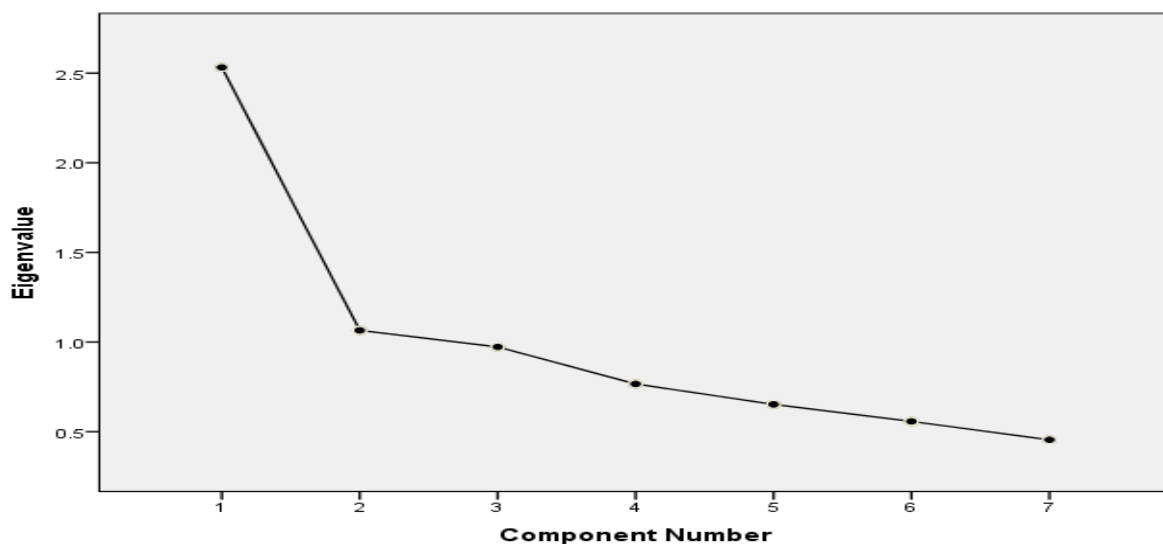


Chart 1. Scree test

Significant main components must be rotated in order to obtain the final factors because the factor analysis is not finished without rotation. Rotations are classified into two groups and they can be orthogonal and oblique. Orthogonal rotations produce factors that are unrelated or independent of each other and oblique rotations produce factors which are interrelated (Suzić, 2007).

Orthogonal factors are easier to interpret, because the interpretation is performed on a single matrix, where after an oblique rotation a pattern matrix, a matrix of structure and a matrix of correlation of factors are obtained (Hasanbegović, 2015). The statistical package SPSS offers varimax, quartimax and equamax rotation that belong to the orthogonal rotations, and direct

oblimin and promax rotations that belong to oblique rotations. The method of varimax rotation tries to minimize the number of variables with high absolute values of factor weights (Pallant, 2007). Out of all orthogonal rotations, varimax rotations are mostly used, and direct oblimin rotations are mostly used out of all oblique rotations. In Table 3 is given an example of the result of direct oblimin rotation, or matrix of form and structure. From the table it can be concluded that out of 7 variables as much were involved in the research (the data is used as an example), the first factor is composed of 5 variables, and the second factor is composed of two variables. After an insight into the structure of the variables that define the factors, it is necessary to give them a title.

Table 3. Parallel and orthogonal projections

Variable	Component		Variable	Component	
	1	2		1	2
VAR00003	.671	↓	VAR00005	.695	↓
VAR00005	.669		VAR00003	.695	
VAR00007	.508		VAR00007	.560	
VAR00001	.485		VAR00002	.535	
VAR00002	.418		VAR00001	.522	
VAR00004		.819	VAR00004		.830
VAR00006		.763	VAR00006		.785

The last result that is shown is the correlation between the obtained factors. In this example, there were two and the resulting coefficients show that there is a correlation between the two factors ($r = 0.52$).

Table 4. Correlation between factors

Factor	1	2
First factor	1	.52
Second factor	.52	1

CONCLUSION

Based on theoretical considerations and practical elaboration, efforts were made in context of acquiring knowledge to point to the possibility and the importance of the application of factor analysis in the field of education and rehabilitation. Considering that the biopsychosocial approach is fundamental for education and rehabilitation, factor analysis can be a powerful tool when studying interconnection of different phenomena. Its proper application by special educators-rehabilitators who deal with these problems, may help in understanding the causes of connections between phenomena, and as such it helps in the development of a treatment for prevention, education and rehabilitation of persons with disabilities.

REFERENCES

- Cattell, R.B. (1966). *The meaning and strategic use of factor analysis*. In R.B. Cattell (ed.), *Handbook of Multivariate Experimental Psychology*. Chicago: Rand McNally.
- Doležal, D. (2007). Faktorska analiza upitnika „Psihološki inventar kriminalnih stilova razmišljanja“. *Hrvatska revija za rehabilitacijska istraživanja*, 43 (1): 13-29.
- Fulgosi, A. (1988). *Faktorska analiza – 3. dopunjeno izdanje*. Zagreb: Školska knjiga.
- Fulgosi, A., Marković, S. (1976). Prilog proučavanju strukture psiholinguističkih sposobnosti-faktorski test jednog psiholinguističkog modela. *Revija za psihologiju*, 6 (1): 67.
- Fulgosi, A. (1979). *Faktorska analiza*. Zagreb: Školska knjiga.
- Hasanbegović, H. (2015). *Osnovi naučnog istraživanja*. Tuzla: Institut za humanu rehabilitaciju.
- Hasanbegović, H., Mehemedinović, S., Mahmutović, E. (2012). Latent structure of motor abilities and skills of deaf children. *Journal for Interdisciplinary Studies „Human“*, 2 (1): 31-36.
- Mejovšek, M. (2003). *Uvod u metode znanstvenog istraživanja*. Jastrebarsko: Naklada Slap.
- Nunnally, J.O. (1978). *Psychometric theory*. New York: McGraw-Hill.
- Pallant, J. (2011). *SPSS*. Beograd: Mikro knjiga.
- Suzić, N. (2007). *Primjenjena pedagoška metodologija*. Banja Luka: XBS.
- Tabachnick, B.G., Fidell, L.S. (2007). *Using multivariate statistics* (5th edn). Boston: Pearson Education.