Australian Journal of Basic and Applied Sciences, 10(18) December 2016, Pages: 143-151

AUSTRALIAN JOURNAL OF BASIC AND APPLIED SCIENCES

ISSN:1991-8178 EISSN: 2309-8414
Journal home page: www.ajbasweb.com

Comparison of Factor Score Computation Methods In Factor Analysis

[1]Müge Uluman and [2]C. Deha Doğan

[1]Marmara University Atatürk Educational Faculty İstanbul - Turkey
[2]Ankara University Educational Faculty Ankara – Turkey

Address For Correspondence:
Müge Uluman, Marmara University Atatürk Educational Faculty İstanbul – Turkey

ARTICLE INFO

Article history:
Received 19 September 2016
Accepted 10 December 2016
Published 31 December 2016

Keywords:
Factor scores, refined and non-refined factor computation methods.

ABSTRACT

Background: Factor scores might be the most useful outcomes of factor analysis. Researchers have to choose one from a variety of methods ranging between simple and sophisticated. Objevtive: Aim of this study is to test whether the use of different factor score computation methods affects the results of research. Methodology:This is a basic research which aims at describing and interpreting the selected case by compiling information. Factor scores were derived with refined and non-refined factor computation methods in real data set. The data were collected from 416 undergraduate students attending Ankara University Educational Sciences Faculty and Marmara University Atatürk Education Faculty. This study examined the correlation coefficients among the factor scores obtained using different factor computation methods for each factor. Result: It was determined that there are high positive correlations among them. When non-refined methods are used the correlations between factor 1 and factor 2 scores were found to be zero. So when refined methods are chosen, the usage of orthogonal rotation is advised since it requires no correlation between the sub-factors. T test results yielded significant difference by gender for the factor scores obtained with all factor score computation methods. ANOVA results revealed that factor scores obtained with all methods other than the sum score method are varied significantly by the departments of the participants. Therefore, all the variables that have an impact on factor scores from the beginning to the end of the factor analysis should be considered so that follow-up analyses can be robust and interpretable.

INTRODUCTION

Researchers who work in social, educational and psychological sciences often need to measure variables (e.g., attitude, perception) that cannot be measured directly. With the development of computer technology and statistical packages, multivariate methods have become prominent and available for researchers. One of the multivariate methods is exploratory factor analysis (EFA). Exploratory factor analysis is a widely used interdependence technique in social, educational and psychological sciences. The essential purpose of EFA is to identify the underlying structure of variables (Hair, Black, Babin, Anderson, & Tatham, 2006). Also, this technique serves different kinds of purposes: constructing a questionnaire or survey to measure latent variables; reducing complex sets of data or identifying groups of variables. Another purpose is using factor information for subsequent studies (Gorsuch, 1983; Adeeb, ALRahamneh & Hawamdeh, 2016). In other words, researchers need scores that represent each individual building up factors for follow-up studies (DiStefano, Zhu, & Mindrila, 2009).

Factor scores might be the most useful outcomes of factor analysis (Tabachnick & Fidell, 2013). There is a variety of subsequent analyses that use factor scores such as multiple regression analyses, analyses of variance, or composite scores. Thus, factor scores perform instead of original data in any subsequent analyses. For

To Cite This Article: Müge Uluman and C. Deha Doğan., Comparison of Factor Score Computation Methods In Factor Analysis. ***Aust. J. Basic & Appl. Sci.,*** *10(18): 143-151, 2016*

example, a researcher administers a t test to examine whether female or male students have more anxieties using the factor scores. Another important use of factor scores is in reducing multi-collinearity. At this point, the utilisation of a factor score rather than the original data as a predictor variable in multiple regression analyses would be more appropriate. Consequently, factor scores are common in social, educational and psychological sciences.

The process of calculating factor scores is not straightforward (Grice, 2001). Researchers have to choose one from a variety of methods ranging between simple and sophisticated (Tabachnick & Fidell, 2013). For example, let us suppose that three researchers carry out an EFA on a questionnaire. Each of them specifies the same number of factors, uses the same number of extractions and rotations. However, each decided to calculate the factor score by a different method. The first one selects the default option which is concluded continuous factor scores (means equal to zero and standard deviations equal to 1) from any computer program. The second researcher chooses the structure coefficients that are the correlations between the salient items and factors. The last researcher uses a summing procedure by selecting the salient items from the factor score coefficient matrix. All the researchers would get different sets of factor scores. Also, they would get different rankings of individuals on the factors (Grice, 2001). The question is 'Which factor score is the appropriate one?' or 'Which factor scores methods are more correct?' This question leads us to debate the indeterminacy of factor scores, and to examine the properties of factor scores methods.

In the literature, lots of factor score indeterminacy discussions have already been done (Mulaik, 1976; Maraun, 1996; Green, 1976; Grice, 2001; Tucker, 1971; Meyer, 1973). Briefly, the indeterminacy problem relates to the creation of an infinite number of sets of factor scores for the same analyses. It is not related to factor scores that cannot be calculated directly (Grice, 2001). The degree of indeterminacy would not be the same in all studies. The reason for this is that indeterminacy depends on the number of items and factors in a specific study (Meyer, 1973). Also, the degree of indeterminacy is related to the factor scores calculation method.

There are two main categories of factor scores calculation called refined and non-refined. Non-refined methods are relatively easy in terms of calculation and interpretation compared to refined methods. Both of them include their own subcategories.

Non-Refined Methods:

The methods that are efficient alternatives for researchers are defined as simple cumulative schemes (Grice, 2001). A wide variety of non-refined methods exist in the literature. However, four commonly used methods are explained here.

Sum Scores by Factor:

This method is probably the simplest one; adding together scores on all items that have a loading on the factor for estimating each individual's factor scores. But if any item has negative loading on the factor then Item's raw score is subtracted instead of added in calculation process because that item is negatively associated with relevant factor. In this method, items which have a greater standard deviation make a greater contribution to the factor scores (Tabachnick & Fidell, 2013). Summed scale items are more appropriate when generalisability is preferred. This method is probably the best one if the summed scale is a well-constructed, valid, and reliable instrument (Hair *et al*., 2013). This method is quite efficient and easy to use although it has some drawbacks. One of those is the equal weight problem. This means that, without considering the loading, all items give equal weight to the relevant factor. In other words, the item that has the highest loading and the item that has the lowest loading have equal weight factor scores. This problem may result in less reliable factor scores. Overleaping the items that are on different metrics or the items that have a different quantity of variability is again related to the equal weight problem (DiStefano, Zhu, & Mindrila, 2009).

Sum Scores – Above a Cut-off Value:

Another simple method is adding up items that are above the cut-off value. Firstly, the researcher has to decide on an arbitrary value which will be a cut-off. Later, items are selected that are above the cut-off value and are added up together to estimate the individual's factor score on the related factor. If any item has a negative loading on the related factor, then the researcher subtracted instead of added the raw score, because this item is negatively related to the factor. The arbitrary selection of a cut-off value is a disadvantage of this method. If the researcher decides to choose a higher cut-off value, then fewer items are used in the factor score estimating process. On the other hand, choosing a lower cut-off value causes fewer related items to be used in the process. Another disadvantage is the equal weight problem, as in the previous method. If variability does not change much among items, then this method's usage might be appropriate.

Sum Scores – Standardised Variables:

This method is more sophisticated than the previous two methods. It includes adding together the standardised items. Firstly, all items are scaled to the same mean and standard deviation. Secondly, we add all standardised items scores loaded on a factor or the items that are above the chosen cut-off value. The choice of process in the second step depends on the researchers. If the standard deviation of the raw data varies in a wide range, the use of this method is more appropriate. If not, this method does not have any advantages over previous methods. In any case, potentially, it still has an equal weight problem (DiStefano, Zhu, & Mindrila, 2009).

Weighted Sum Scores:

The first three methods ignore the items weights which base their loadings on the factor. This means that items with lower loadings on the factor have the same weight as the items with higher loadings. So all items have the same weight when the factor scores are computed. However, items with higher loadings might have a larger effect on the total factor score and vice versa. In this method, the weights vary across items in accordance with the factor loading of that item on a related factor. Before the addition, the factor loading of each item is multiplied to the scale score for each item (DiStefano, Zhu, & Mindrila, 2009). The summing process can be applied just to items which are above a specific cut-off value or to all items. That decision belongs to the researchers.

This method's advantage is that it allows those variables with the highest loadings on the factor to have the greatest effect in estimating the factor scores. However, one of the potential problems with this method is that the factor loadings may not be an accurate representation of the differences among factors due to a researcher's choice of extraction model and/or rotation method (DiStefano, Zhu, & Mindrila, 2009). To the extent that this is true, this method would not represent an improvement over the previous one.

The advantage of non-refined methods is their stability across independent samples of observation relative to refined methods (Grice & Harris, 1998). This means that the obtained results do not depend heavily on the particular sample used (DiStefano, Zhu, & Mindrila, 2009). Another advantage is that these methods, with the use of statistical software, calculate the factor scores very easily. However, non-refined methods suffer from a number of defects. For instance, they may be highly correlated even when the factors are orthogonal and they will be less valid representations of the factors in comparison with refined methods.

Refined Methods:

Several other methods than the non-refined methods have also been developed that provide linear approximations of the factor scores using a complex refined W_{kf} weight matrix (Grice, 2001). Refined methods' factor scores are linear combinations of the observed variables which consider what is shared between the item and the factor and what is not measured (Gorsuch, 1983;DiStefano, Zhu, & Mindrila, 2009).

Methods in this category aim to maximise validity by producing factor scores that are highly correlated with a given factor and to obtain unbiased estimates of the true factor scores (DiStefano, Zhu, & Mindrila, 2009). The most common refined methods are defined here.

Regression Scores:

The most common refined methods use standardised information to create factor scores, producing standardised scores similar to a Z-score metric, where values range from approximately -3.0 to +3.0. However, instead of the unit of standard deviation, the exact value can vary (DiStefano, Zhu, & Mindrila, 2009).

Regression methods can also be employed to estimate factor scores. In this method, the factor loadings are adjusted to take account of the initial correlations between variables; in doing so, differences in units of measurement and variable variances are stabilised (Field, 2009). Therefore, this method uses the following basic equation:

$$z_{fi} = \beta_1 z_{1i} + \beta_2 z_{2i} + \beta_3 z_{3i} + \dots + \beta_n z_{ni} \qquad (1)$$

where
z_{fi} is a standard score in factor f for person i,
z_{1i} is a standard score in variable 1 for person i,
z_{2i} is a standard score in variable 2 for person i,
β_i is the standard regression coefficient for variable i.

Equation (1) is like the standard multiple regression equation where n predictors are being used to predict a single criterion variable. To obtain the β_i weight for this equation, it is sufficient to know the correlations among the predictors and the correlation of the predictors with the criterion, that is, the validity coefficients. In the application to the problem of estimating factor scores, the factor scores become the predicted criterion scores, the variables in the factor analysis are predictors, and the orthogonal factor loadings or oblique structure coefficients are the validity coefficients (Grice & Harris, 1998). There are a number of different sets of constraints on the estimated factor scores and attempts to minimise a particular estimate of error. The most popular solution is probably Thurstone's (1935) least squares regression approach in which the factor score coefficients are computed from the original item correlations and structure coefficients (Grice, 2001).

$$\beta_1 + \beta_2 r_{12} + \beta_3 r_{13} + \ldots + \beta_n r_{1n} = r_{1f}$$
$$\beta_1 r_{21} + \beta_2 + \beta_3 r_{23} + \ldots + \beta_n r_{2n} = r_{2f}$$
$$\beta_1 r_{31} + \beta_2 r_{32} + \beta_3 + \ldots + \beta_n r_{3n} = r_{3f}$$
$$\beta_1 r_{n1} + \beta_2 r_{n2} + \beta_3 r_{n3} + \ldots + \beta_n = r_{nf} \quad (2)$$

Equation (2) may be expressed in matrix form as follows:

$$\mathbf{R}\beta = r_f \quad (3)$$

where R is the matrix of known correlations among variables 1 through n in equation (2); β is a column matrix containing the unknown β_i weights; and r_f is a column matrix of correlation between the variables and the factor, that is, orthogonal factor loadings of oblique structure coefficients. Provided the matrix R has an inverse, equation (3) and hence equations (2) may be solved as follows:

$$\beta = \mathbf{R}^{-1} r_f \quad (4)$$

Thus, the column of β_1 weights to be used in equation (1) for predicting the factor scores from the data-variable scores is obtained by multiplying the inverse of the matrix of correlations among the data variables by the column matrix of correlations of the data variables with the factor. An advantage of using all variables in the regression equation (1) is that the inverse $\mathbf{R}^{-1}$ may be obtained once and used in equation (4) for all factors merely by changing the r_f column, depending on which factor is being considered. If only some of the variables are used to obtain a given set of factor scores, the $\mathbf{R}^{-1}$ matrix for predicting that factor must be derived from an **R** matrix containing only those variables that are being used.

Bartlett Scores:

Bartlett (1937) developed a method that minimises the sum of squares for unique factors across the range of variables:

$$W_{kf} = U_{kk}^{-2} P_{kf} (P'_{fk} U_{kk}^{-2} P_{kf})^{-1} \quad (5)$$

Harman (1976) reports the 'idealized variable' strategy based on the reproduced correlation matrix rather than on the original item correlations:

$$W_{kf} = (P_{kf} P'_{fk})^{-1} P_{kf} \quad (6)$$

Correlations between the factors and non-corresponding factor score estimates computed from the weight matrices in equations 5 and 6 are constrained to 0 when the factors are orthogonal (Grice, 2001). With the Barlett approach, only the shared factors have an impact on factor scores. The sum of squared components for the 'error' factors across the set of variables is minimised, and resulting factor scores are highly correlated to their corresponding factor and not with other factors. However, the estimated factor scores between different factors may still correlate (DiStefano, Zhu, & Mindrila, 2009).

Anderson-Rubin Scores:

An additional linear prediction method called the Anderson and Rubin method was developed as an alternative to the regression method. Anderson and Rubin (1956) developed a procedure for estimating factor scores that are constrained to orthogonality:

$$W_{kf} = U_{kk}^{-2} P_{kf} (P'_{fk} U_{kk}^{-2} R_{kk} U_{kk}^{-2} P_{kf})^{-1/2} \quad (7)$$

where P_{kf} represents the pattern coefficients and U_{kk} is a diagonal matrix of the reciprocals of the squared unique factor weights. This equation is appropriate even when the covariance matrix for the unique factors is non-singular, although in practice this matrix is assumed to be diagonal (Tucker, 1971; Grice, 2001). In this method the least squares formula is adjusted to produce factor scores that are not only uncorrelated with other factors, but also uncorrelated with each other. The resulting factor scores are orthogonal, with a mean of 0 and a standard deviation of 1 (DiStefano, Zhu, & Mindrila, 2009).

As can be seen, the factor score computation methods presented for the use of researchers have different advantages and disadvantages relative to each other. Still, researchers have no clear information or criteria that may guide the selection of the most appropriate of these methods. In addition, the literature does not provide adequate studies investigating whether the use of different methods affects research results.

The first aim of this study is to draw attention to the presence of different factor score computation methods that allow for working with a smaller number of variables in areas of education, psychology, and social sciences which often use factor analysis. The results of the studies carried out by using the scores obtained with factor analysis are shaped in accordance with these scores. If the points calculated for the same factor vary as a consequence of using different methods, the study results may also change. Therefore, it is an important to know whether the use of these methods affects the results of research. It is also of high importance that the researcher knows the advantages and disadvantages of the factor analysis method to be used and that s/he selects the most appropriate method for the purpose of the research.

Australian Journal of Basic and Applied Sciences, 10(18) December 2016, Pages: 143-151

In this context, factor scores were derived with refined and non-refined factor computation methods in the real data set investigated in the study. These points were used to seek responses to the following sub problems of this study:

1. What is the correlation between the factor scores of the "Study Process Questionnaire" computed using different factor score computation methods?
2. What is the correlation between the factor 1 and the factor 2 scores of the "Study Process Questionnaire" when different factor score computation methods are used?
3. Does the use of different factor score computation methods affect the results of the statistical tests used to compare mean scores according to categorical independent variables with two or more subgroups?

a. Does the use of different factor score computation methods affect the t test results according to gender?

b. Does the use of different factor score computation methods affect the one-way ANOVA results according to class level?

Methodology:

Research Model:

This study was carried out to compare different methods used in the factor creation process from various aspects, and it was carried out in a basic research model. Basic research aims at describing and interpreting the selected case by compiling information. In this kind of research, the intention is to understand the issue in the best way possible, to try to complete the information, or to develop a new theory (Wallen & Fraenkel, 2001).

Study Group:

The study data were collected from 416 undergraduate students attending Ankara University Educational Sciences Faculty and Marmara University Atatürk Education Faculty. The study aim was to compare and contrast different methods in the light of the data collected from the participants beyond providing a description of the study group. Thus, the convenience sampling method was used. The demographic information on the participants is given in Table 1 including their university, gender and grade levels.

Data Collection Instruments:

In this study, the data were collected by using the "Study Process Questionnaire", which was developed by Biggs, Kember and Leung (2001) and adapted to the Turkish culture by Beşoluk and Önder (2010). As a result of the Exploratory Factor Analysis calculated during the adaptation study of the scale, it was determined that the scale yielded a two-factor structure, as in the original version, and explained approximately 35% of the total variance. Then, the resulting two-factor structure was confirmed with the confirmatory factor analysis. The resulting fit indices ($\chi 2/sd$) = 2.94, RMSEA = .061, NFI = .90, CFI = .93, IFI = .93, RFI = .88, GFI = .92, AGFI = .89 and SRMR = .065) were found to indicate medium and high levels of fit.

In the study, the scale should have more than one 1 factor in order to compare the factor score computation methods. However, we used a scale with a 2-factor structure for a more practical comparison of the factor score computation methods. The "Study Process Questionnaire" was used as the data collection tool since it has a 2-factor structure, it does not contain too many items allowing practical application, it can be given to university students, and it has the required level of validity and reliability.

The purpose of this study is not to determine the psychometric properties of the scale concerned or to describe the group to which the scale was given. Therefore, the data obtained from the data collection tool were only used in the exploratory factor analysis, and to compare different methods used for determining the factor scores in the factor analysis process.

Data Analysis:

First, an exploratory factor analysis was performed on the study data. Then, different factor scores were calculated for the obtained factors by using "total factor scores", "sum scores above a cut-off value", "sum scores for standardised variables", "weighted sum scores", "regression scores", and "Bartlett scores" methods.

After that, the relationship between the sequence of individuals was calculated by using the Spearman Brown sequence differences correlation coefficients for factor scores obtained with different methods. Then, Fisher's z formula was applied to find out whether such differences are at a significant level. In addition, the relationship between the factor score 1 and 2 obtained with different methods was tested by using the Pearson correlation coefficient, and again the significance of the differences between correlation coefficients was tested with Fisher's z formula.

A t test for independent samples and a one-way variance analysis were conducted to test whether the scores in both methods differ by gender and grade level on the basis of the participants' factor scores calculated in different ways. The assumptions of homogeneity of univariate normality and variances were examined prior to calculating the respective hypotheses.

Findings:

Correlation among Factor Score Computation Methods:

The Spearman Brown Correlation Coefficients were calculated among different factor computation methods in order to see whether the ranking of the participants who answered the "Study Process Questionnaire" differs according to the methods used to calculate the factor scores. Table 1 shows the correlation coefficients.

Table 1: The Spearman Brown Correlation Coefficients among factor computation methods

Factor Score Computation Methods		Sum Scores	SS-Above a cut-off Value	SS-Standardised Variables	Weighted sum scores	Regression Scores	Bartlett Scores	Anderson-Rubin Scores
Non-Refined Methods	Sum scores	1.000						
	Sum scores above a cut-off value	.964**	1.000					
	Sum scores for standardised variables	.954**	.991**	1.000				
	Weighted sum scores	.958**	.992**	.977**	1.000			
Refined Methods	Regression scores	.943**	.965**	.972**	.961**	1.000		
	Bartlett scores	.943**	.965**	.972**	.961**	1.000	1.000	
	Anderson-Rubin scores	.943**	.965**	.972**	.961**	1.000	1.000	1.000

As it is seen in table 1, all correlations coefficients among different factor score manipulations methods are positive, strong, and significant at 001 level. The minimum correlation coefficients (0.943) are between the sum scores method and the refined methods (Regression scores, Bartlett Scores, Anderson-Rubin Scores). The maximum correlation coefficients (1.00) are among the regression scores, Bartlett Scores, and Anderson-Rubin Scores methods. Because the correlation coefficients are very strong, positive and significant, the ranking of the participants who answered the "Study Process Questionnaire" does not differ according to the methods used to calculate factor scores. So either of them may be used if it is aimed to rank participants according to factor scores.

Correlations between Factor 1 and Factor 2 Scores according to Different Factor Score Computation Methods:

The Spearman Brown Correlation Coefficients were calculated between factor 1 and factor 2 scores for each factor score computation method separately. The aim was to see whether the ranking of the participants who answered the "Study Process questionnaire" differs according to factor 1 and factor 2 scores for different factor score computation methods. Table 2 summarises the correlation coefficients between factor 1 and factor 2 scores for different factor score computation methods.

Table 2: Correlation between factor 1 and factor 2 scores for different factor score computation methods

Factor Score Computation Methods		Correlation between Factor 1 and Factor 2 Scores
Non- Refined Methods	Sum scores	-0.331**
	Sum scores above a cut-off value	-0.332**
	Sum scores for standardised variables	-0.327**
	Weighted sum scores	-0.346**
Refined Methods	Regression scores	.000
	Bartlett scores	.000
	Anderson-Rubin scores	.000

As it is seen in table 2, when non-refined methods (sum scores, sum scores above a cut-off value, sum scores for standardised variables, weighted sum scores) are used to compute factor scores, there are negative and moderately strong correlations between factor 1 and factor 2 scores.

Those correlation coefficients are significant at 0.01 level. For non-refined methods, the strongest correlation coefficient (-0.35) was calculated when weighted sum scores were used and the weakest correlation coefficient (-0.33) was calculated when sum scores for standardised variables were used. Fisher's z test was performed to test the differences between correlation coefficients for non-refined models and non-significant results were found. So, using any of the non-refined methods gives similar results if the aim is to test the correlation between the sub-factors of a scale.

On the other hand, the results for refined methods (regression scores, Bartlett scores and Anderson-Rubin scores) are strictly different from the ones for non-refined models. Table 2 shows that when non-refined methods are used to compute factor scores, the correlations between factor 1 and factor 2 scores were found to be zero. It can be said that non-refined models ignore the correlation between sub factors and equalise it to zero. So if the aim is to test the correlation between the sub-scores of a scale, the use of refined methods is not advised since it ignores the correlation and increases the type II error. Moreover, while calculating factor analysis, when refined methods are used, the usage of orthogonal rotation is advised since it requires no correlation between the sub-factors of the scale.

Comparison of Group Means for Different Factor Score Computation Methods:

An independent sample t test and a one-way ANOVA were calculated to compare significance of group means for each factor score computation method. First, the independent sample t test was calculated to test if the mean score of female and male participants differs significantly. This calculation was repeated for each factor score computation method and t scores, p values and effect sizes were compared. Table 3 shows the independent sample t test results for each factor score computation methods separately.

Table 3: Independent sample t test results for various factor computation methods

Factor Score Computation Methods		t score	P value	Effect size
Non- Refined Methods	Sum scores	-3.895	.000	0.187
	Sum scores above a cut-off value	-3.345	.001	0.161
	Sum scores for standardised variables	-3.091	.002	0.148
	Weighted sum scores	-3.688	.000	0.176
Refined Methods	Regression scores	-2.718	.007	0.130
	Bartlett scores	-2.718	.007	0.130
	Anderson-Rubin scores	-2.718	.007	0.130

Table 3 shows that although t scores differ among factor score computation methods, all of them have similar p values which are significant at 0.001 level. So, it can be said that the significance of the difference between two independent variables does not change according to the factor score computation methods. Moreover, effect sizes were calculated for each factor score computation method separately. Results show that total sum scores method has the highest effect size (0.19) and that the refined methods (Regression scores, Bartlett scores, Anderson-Rubin scores) have the lowest effect sizes (0.13). Fisher's z test was calculated to test the significance of the differences between effect sizes and no significant results were found. So, all the effect sizes calculated for different factor score computation methods have a medium effect. In other words, when an independent sample t test is calculated, the effect size does not differ according to the factor score computation method.

Subsequently, a one-way analysis of variance (One Way ANOVA) was calculated to test if the mean score of participants from various classes (2, 3, 4 and 5) differs significantly. This calculation was repeated for each factor score computation method and F values, p values and effect sizes were compared. Table 4 shows the one-way ANOVA results for each factor score computation method separately.

Table 4: One-way ANOVA results for various factor computation methods

Factor Score Computation Methods		F value	P value	Effect size
	Sum scores	2.578	.053	0.018
Non- Refined Methods	Sum scores above a cut-off value	2.578	.042*	0.019
	Sum scores for standardised variables	2.673	.047*	0.019
	Weighted sum scores	2.900	.035*	0.021
Refined Methods	Regression scores	3.493	.016**	0.024
	Bartlett scores	3.493	.016**	0.024
	Anderson-Rubin scores	3.493	.016**	0.024

As it is seen in Table 4, the significance of F value differs according to the factor score computation methods. When refined methods (regression scores, Bartlett score, Anderson-Rubin score), sum scores Above a cut-off value method, sum scores for standardised variables method and weighted sum scores methods are used to compute factor scores, F values were significant at the 0.05 level. On the other hand, when the sum scores method was used, F value was not statistically significant. So using sum score methods to compute factor scores may increase type II error while calculating one-way ANOVA. However, when effect sizes are examined, it can be seen that all of them have a very low value which shows little effect. So, although all other methods except the sum score method show significant F value, they have very low effect size. So, the calculation of effect size is very important when using different factor score computation methods if the aim is to calculate one-way ANOVA.

Conclusion:

Factor analysis is commonly used in education, psychology and social sciences. If new analyses are to be performed based on the results obtained, the choice of the method of calculating factor scores becomes important.

The selection of a factor score computation method might differ depending on the aim of the study and the observed correlation coefficient among factors. Though there are not strict rules to confirm the accuracy of the selection, there are certain indicators. In the scope of the study, these indicators were investigated and recommendations were presented to researchers.

These methods were investigated in two categories (non-refined and refined methods). Non-refined calculation methods are quite easy and stable, but the validity and accuracy is low. However, refined calculation methods can be said to be more difficult and complex, but more valid. Nevertheless, they can be calculated quite easily with developed statistical software packages.

This study examined the correlation coefficients among the factor scores obtained using different factor calculation methods for each factor and it was determined that there are high positive correlations among them.

While the correlation between the first and second factor results in significant difference between non-refined and refined methods, the non-refined and refined methods do not pose a significant difference within themselves. When non-refined methods are used to compute factor scores, the correlations between factor 1 and factor 2 scores were found to be zero. It can be said that non-refined models ignore the correlation between sub factors and equalize it to zero. According to DiStefano, Zhu and Mindrila, (2009) refined methods attempt to retain the relationships between factors. Because factor scores are based on factors and the relationship among items, obliquely or orthogonally transformations may also lead to differences in the resulting factor scores (Grice, 2001).

Similarly, factor score calculation methods can influence the fitness of the selected transformation in the context of the study carried out. So, while calculating factor analysis, when refined methods are chosen, the usage of orthogonal rotation is advised since it requires no correlation between the sub-factors of the scale.

Considering gender and department variables in this study, t test and one-way analysis of variance were calculated for the factor score obtained with each of the factor score calculation methods. T test results yielded significant difference by gender for the factor scores obtained with all factor score computation methods. The effect sizes calculated for each method were also found to be similar. One-way ANOVA results revealed that factor scores obtained with all methods other than the sum score method are varied significantly by the departments of the participants. However, when effect sizes are examined, it can be seen that all of them have very low value which shows a little effect. So, although other methods except the sum score method show significant F value, they have very low effect size. So, the calculation of effect size is very important when using different factor score computation methods if the aim is to calculate the one-way ANOVA.

On the other hand, in the selection of the factor score calculation method, the fit with the different variables used should be considered as well as the purpose of the follow-up analysis. As stated by Zuccaro (2007), if one of the variables used contains standardised values, it would be more appropriate that factor scores are also based on standardised methods.

It should also be noted that regardless of the method used, factor scores can be affected by the factor extraction method performed during EFA analysis (DiStefano, Zhu, & Mindrila, 2009). Therefore, all the variables that have an impact on factor scores from the beginning to the end of the factor analysis should be considered so that follow-up analyses can be robust and interpretable.

REFERENCES

Adeeb A. Ali ALRahamneh and Omar M. Hawamdeh, 2016. Using Factor Analysis to Analyze Slight and Sever Injury by Age Group and Road User in Jordan. Australian Journal Of Basic And Applied Sciences, 10(6): 86-98.

Anderson, R.D. and H. Rubin, 1956. Statistical inference in factor analysis. Proceedings of the Third Berkeley Symposium of Mathematical Statistics and Probability, (5): 111-150.

Bartlett, M.S., 1937. Properties of sufficiency and statistical tests. Proc of the Royal Statistical Society, Series A, 160: 268-28.

Beşoluk, Ş. and İ. Önder, 2010. Öğretmen adaylarının öğrenme yaklaşımları, öğrenme stilleri ve eleştirel düşünme eğilimlerinin incelenmesi. İlköğretim Online, 9(2): 769-693.

Biggs, J., D. Kember and D.Y. Leung, 2001. The revised two-factor study process questionnaire: R-SPQ-2F. British Journal of Educational Psychology, 71(1): 133-149.

DiStefano, C., M. Zhu and D. Mindrila, 2009. Understanding and using factor scores: Considerations for the applied researcher. Practical Assessment, Research & Evaluation, 14(20): 1-11.

Field, A., 2005. Discoveringstatistics using SPSS (2. bs). London:Sage Publications.

Green Jr, B.F., 1976. On the factor score controversy. Psychometrika, 41(2): 263-266.

Grice, J.W., 2001. A comparison of factor scores under conditions of factor obliquity. Psychological methods, 6(1): 67-83

Grice, J.W. and R.J. Harris, 1998. A comparison of regression and loading weights for the computation of factor scores. Multivariate Behavioral Research, 33(2): 221-247.

Gorsuch, R., 1983. *Factor analysis*. Hillsdale, NJ: L. Erlbaum Associates.

Hair, J.F., W.C. Black, B.J. Babin, R.E. Anderson and R.L. Tatham, 2006. *Multivariate data analysis* (Vol. 6). Upper Saddle River, NJ: Pearson Prentice Hall.

Maraun, M.D., 1996. Metaphor taken as math: indeterminancy in the factor analysis model. Multivariate Behavioral Research, 31(4): 517-538.

Meyer, E.P., 1973. On the relationship between ratio of number of variables to number of factors and factorial determinacy. Psychometrika, 38(3): 375-378.

Mulaik, S.A., 1976. Comments on "The measurement of factoriindeterminacy". Psychometrika, 41(2): 249-262.

Tabachnick, B.G. and L.S. Fidell, 2001. Using Multivariate Statistics (4th Ed.). Needham Heights, MA: Allyn & Bacon.

Thurstone, L..L., 1935. The vectors of mind. Chicago: University of Chicago Press., pp: 226-231.

Tucker, L.R., 1971. Relations of factor score estimates to their use. Psychometrika, 36(4): 427-436.

Wallen, N.E., J.R. Fraenkel, 2001. Educational research: A guide to the process (2nd ed.). Mahwah, NJ: Lawrence Erlbaum Associates.

Zuccaro, C., 2007. Statistical alchemy--The use of factor scores in linear regression. Retrieved september 6, 2007, from http://www.esg.uqam.ca/recherche/document/2007/ 01-2007.pdf