IOSR Journal of Mathematics (IOSR-JM)
e-ISSN: 2278-5728, p-ISSN: 2319-765X. Volume 15, Issue 5 Ser. II (sep – Oct 2019), PP 33-40
www.iosrjournals.org

Exploratory Factor analysis: Domestic Violence on Women In Saudi Arabia

Hadeel F. Alrowithy[1], Salwa L. AlKhayyat[2] and Ali H. ALgarni[3]
[1]Department of Statistics, Sciences Faculty, Jeddah of University, Jeddah, Saudi Arabia
[2]Department of Statistics, Sciences Faculty, Jeddah of University, Jeddah, Saudi Arabia
[3]Department of Statistics, Faculty of Science, King Abdulaziz University, Jeddah, Saudi Arabia

Abstract: *Violence against women is one of the most problematic issues that hinder the development and growth of societies. This researcher paper aims to deciding the most important reasons that lead to violence against women in the context of Saudi Arabia. This objective is achieved by using exploratory factor analysis which is used to explain the relationship among the variables and to decrease the number of variables. This factor analysis is applied on a sample that includes 800 women from different regions of Saudi Arabia. The study consists of 18 variables. Correlation matrix is checked and the extraction method of Principal components analysis (PCA) is applied by using varimax orthogonal rotation, we get 6 factors that explain around 60 % from total variance respectively, the work factor, the family factor, the frequency of committing violence,the abuser factor, the income and the surroundings factor.*

Keywords: *Domestic violence, Factor Analysis, Scree Test, Extraction, Eigenvalues*

Date of Submission: 11-09-2019 Date of Acceptance: 26-09-2019

I. Introduction

Women play an important role in strengthening economic, social, cultural and political issues therefore, Saudi Vision 2030 aims to increase the role of Saudi women and their participation in the labor market and invest their capacities in the development of the country. However, many women face various challenges. One of the most important challenges is the violence they face at home, which is called Violence against women (VAW). It is also known as gender-based violence [1].

United Nation defines violence against women as "any act of gender-based violence that results in, or is likely to result in, physical, sexual or psychological harm or suffering to women, including threats of such acts, coercion or arbitrary deprivation of liberty, whether occurring in public or in private life". [2]

The violence limits women's contribution to the society development and reduce their productivity. There are many factors that lead to the committing violence within the family, whether psychologically or physically. Consequently, it affects the performance of women inside and outside their homes and it is reflected on the community.

In this paper, we will study the violence against Saudi women to determine the significant factors of violence against Saudi woman and to make recommendations, this will help reduce violence and address problems that are caused by violence.

To achieve this objective, the factor analysis will be applied. Factor analysis is defined as a procedure of a multivariate statistical which is used in many ways in order to decrease the used variables to be smaller, build the abstract dimensions between latent and measured variables. [3]

Factor analysis has two main classes; Confirmatory Factor Analysis (CFA) and Exploratory Factor analysis (EFA). That helps explore the most major dimensions in order to produce a theory, or to generate a model from a greater latent construct set which is often offered by a group of items. However, the researcher in CFA, a set of structured equation modelling make use of that approach in order to examine a suggested theory.
It also examines a model which is contrasted to EFA that has expectation and assumptions which are based on preview theory. Concerning a set of factors in order to find out the best factor theories and models can fit. [4]

In this study, the reasons that to violence will be determined by using EFA. Exploratory factor analysis (EFA) can be defined as the mothed of a multivariate statistic which is turned to be a necessary means in the psychological measurements and theories of validation and development. [5]

II. Material and Methods

In this section, the concept of factor analysis and its application will be illustrated using SPSS, an application to the data of Saudi women who have been subjected to violence.

The required data is taken from a questionnaire, the data included 800 Saudi women from different regions of the Kingdom of Saudi Arabia the period from December 2018 to August 2019 and 18 variables:
Q_1: Age. Q_2:Marital status.Q_3: educational level,
Q_4: number of family members.Q_5: place of residence.Q_6: type of house,
Q_7: Administrative Area.Q_8 : Work.Q_9: Work Type,
Q_{10}: Working Hours.Q_{11}: Income.Q_{12} : Age of Violence,
Q_{13} : Times of Violence.Q_{14}: Type of Violence.Q_{15} : Abuser,
Q_{16} : Age of abuser.Q_{17}: Job of abuser and Q_{18} : Income of abuser.

Exploratory Factor Analysis is a multi-step process, which can be classified into three basic steps:
1- The suitability of data for factor analysis.
2- Extraction method(The initial solution).
3- Rotation method (The final solution).

In this paper, the steps are explained in addition to their applications to the data of the study using the statistical program SPSS which is one of the most popular applications used in psychological research and social science.

2.1 The suitability of data for factor analysis:

Factor analysis stems from the matrix of correlation,the correlation matrix the variables should be intercorrelated; however, they do not have to correlate very highly. there would be no correlations between the variables. Multicollinearity, then, can be detected via the determinant of the correlation matrix, if the determinant is greater than 0.00001, then there is no multicollinearity.[6]

Table (1): Correlation Matrix

		Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9	Q10	Q11	Q12	Q13	Q14	Q15	Q16	Q17	Q18
Correlation	Q1	1.000	.466	.091	-.139	.333	.117	-.020	.376	.311	.311	.118	.428	-.006	.065	.389	.151	.079	.094
	Q2	.466	1.000	-.033	-.274	.249	-.049	-.014	.221	.223	.206	.001	.231	.046	-.040	.344	-.121	.068	-.004
	Q3	.091	-.033	1.000	-.151	.016	.151	.032	.353	.282	.299	.066	.080	-.087	.028	.065	.026	.077	.034
	Q4	-.139	-.274	-.151	1.000	-.314	.078	.013	-.207	-.130	-.150	-.074	-.040	.075	-.030	-.239	.182	-.121	.049
	Q5	.333	.249	.016	-.314	1.000	-.026	-.054	.128	.081	.081	.141	.115	-.061	.131	.459	-.110	.226	.035
	Q6	.117	-.049	.151	.078	-.026	1.000	.193	.139	.084	.124	-.010	.050	-.024	.052	.002	.101	.121	.131
	Q7	-.020	-.014	.032	.013	-.054	.193	1.000	-.030	-.082	-.042	-.047	-.027	.047	-.012	-.061	.057	.027	.016
	Q8	.376	.221	.353	-.207	.128	.139	-.030	1.00	.780	.866	-.048	.154	-.047	.060	.175	.014	.077	.066
	Q9	.311	.223	.282	-.130	.081	.084	-.082	.780	1.00	.772	-.034	.145	-.024	.026	.155	-.020	.005	.061
	Q10	.311	.206	.299	-.150	.081	.124	-.042	.866	.772	1.00	-.057	.145	-.049	.031	.145	-.015	.021	.062
	Q11	.118	.001	.066	-.074	.141	-.010	-.047	-.048	-.034	-.057	1.000	.064	-.035	.034	.139	-.033	.109	.235
	Q12	.428	.231	.080	-.040	.115	.050	-.027	.154	.145	.145	.064	1.000	.219	-.053	.156	.027	-.016	.106
	Q13	-.006	.046	-.087	.075	-.061	-.024	.047	-.047	-.024	-.049	-.035	.219	1.000	-.245	-.045	-.018	-.058	-.003
	Q14	.065	-.040	.028	-.030	.131	.052	-.012	.060	.026	.031	.034	-.053	-.245	1.000	.136	.184	.098	.014
	Q15	.389	.344	.065	-.239	.459	.002	-.061	.175	.155	.145	.139	.156	-.045	.136	1.000	-.120	.185	.080
	Q16	.151	-.121	.026	.182	-.110	.101	.057	.014	-.020	-.015	-.033	.027	-.018	.184	-.120	1.000	-.006	.001
	Q17	.079	.068	.077	-.121	.226	.121	.027	.077	.005	.021	.109	-.016	-.058	.098	.185	-.006	1.000	.056
	Q18	.094	-.004	.034	.049	.035	.131	.016	.066	.061	.062	.235	.106	-.003	.014	.080	.001	.056	1.000
a. Determinant = .010																			

Table (1) represents correlation matrix among variables. It is observed that the correlation matrix is empty from high correlations that exceed 0.8. In spite of the fact that there are some weak correlation coefficient among variables, this variable are necessary in this study and can not be deleted. It is seen that the determinant matrix that exists under the previous table is 0.010 and it is higher than 0.00001. This indicates that there is no multicollinearity among variable. This means that the matrix is not a single one.

The correlation test and determinant is counted not enough for recognizing how suitable is the data to the factor analysis. As a result, we **use Kaiser-Meyer-Olkin (KMO) Measure of Sampling Adequacy and Bartlett's test for Sphericity**, KMO Is defined as a test that is used for assessing if the data is appropriate and suitable to be used for FA. The relationship among the variable is signified through using higher KMO value. If KOM value is more then 0.60, it will show that it is appropriate.[7]Bartlett's test is used for assessing the assumption of the statistical methods can show the variances of population.[8]

Table (2): KMO and Bartlett's Test

Kaiser-Meyer-Olkin Measure of Sampling Adequacy		.739
Bartlett's Test of Sphericity	Approx. Chi-Square	3675.927
	df	153
	Sig.	.000

From table (2) the value of KMO is 0.739 that mean the degree of common variance among variables is middling; this means that the sample is enough to perform FA and Bartlett's Test of Sphericity Statistically significant Where that the P <0.005 ,that mean The sample intercorrelation matrix did not come from a population in which the intercorrelation matrix is an identity matrix. Thus, the data are suited for factor analysis. After checking how suitable is the data to FA, we will process the next step.

2.2: Extraction method

When the variables are decided, and the matrix of correlation is ready, the researcher use factor analysis in order to recognize the latent structure of the different relationships.

There are many methods available for extract factors, such as Principal components analysis (PCA), maximum likelihood (MLF), principal axis factoring (PAF), image factoring (IF), etc. Depending on the research aim. In this paper, we will discuss the Principal components analysis (PCA).

In (PCA), the totality of variables is produced in the components that are resulted of the variance found in original data.[9]

One of the methods that is widely for extraction is(PCA) to derive the maximum variance, factor weights can be used computed. This process will continue till no meaningful variance is left any more.[10]

From the table(3):There are as many components extracted, 18 factors (components) were extracted, the same as the number of variables.

This is not suitable with the goal of FA which is; decreasing the number of the measured variables to a fewer number of latent variables. There are some factors that should be kept.

Some of the methods that can be used to determine the number of factors are:

1- Kaiser Rule:

Kaiser rule is a method that is considered the easiest one as well as the most common method of all. It states that to get all the components with the eigenvalues which are larger than 1.0 procedure. Such a method presents an estimate of an optimal number of the components which are used to depict the given data.[11]

In using PCA every factor that is extracted then it is shown by an eigenvector. An eigenvector is defined as a weight column, each of them accompanies an element in the correlation matrix. The eigenvalue for a given factor measures the variance in all the variables which is accounted for by that factor. The eigenvalues that are used for each factor that is extracted become normally greatest for the first factor that is extracted and the fewest for the last extracted factor.[12]

The eigenvalue for a given factor measures the variance in all the variables which is accounted for by that factor.

Table (3): Extraction Method: Principal Component Analysis.

Total Variance Explained									
	Initial Eigenvalues			Extraction Sums of Squared Loadings			Rotation Sums of Squared Loadings		
Component	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	3.528	19.600	19.600	3.528	19.600	19.600	2.917	16.203	16.203
2	1.973	10.961	30.561	1.973	10.961	30.561	2.484	13.799	30.002
3	1.533	8.516	39.077	1.533	8.516	39.077	1.501	8.341	38.343
4	1.453	8.075	47.151	1.453	8.075	47.151	1.345	7.470	45.813
5	1.211	6.726	53.877	1.211	6.726	53.877	1.323	7.351	53.163
6	1.150	6.392	60.269	1.150	6.392	60.269	1.279	7.105	60.269
7	.908	5.044	65.313						
8	.886	4.925	70.237						
9	.786	4.368	74.605						
10	.751	4.174	78.779						
11	.695	3.859	82.639						
12	.663	3.683	86.322						
13	.655	3.636	89.958						
14	.604	3.354	93.313						
15	.480	2.664	95.977						
16	.356	1.977	97.954						
17	.245	1.359	99.312						
18	.124	.688	100.000						

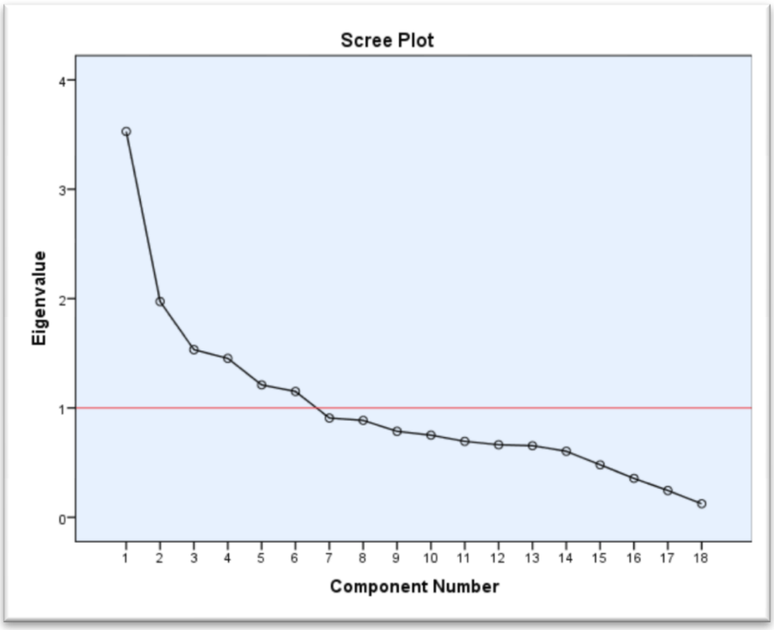
From the table (3):

- From factor 1 to factor 6 have eigenvalues greater than 1, and therefore, which is explains that the morevariance than a single variable. The first factor has its largest eigenvalues3.528 and explains 19.600 % from the total variance. The sixth factor has its smallest eigenvalues 1.150 and explains 6.392 % from the total variance.
- From factor 7 to factor 18 have eigenvalues that are less than1, and therefore, that explains the less variance that a single variable.
- There are six components that explain 60.269% of the common variance between the 18 variables.

2- The Scree Test:

Scree test is a procedure which Cattell (1966) proposed. With scree test, the eigenvalues are used against the numbers of ordinal. One should search for the place of break or the slop level of the plotted line. The pint of break as the point in which a line is drawn through and the points can change it is direction. The number of eigenvalues indicates the number of components which occur above the break point.[13]The error variance is indicated by the eigenvalues that are below the break pint.[14]

Figure (1): Scree Plot



From Figure (1)the number of the six extracted factors appears above the point from which the carve starts to slow down.
The scree plot is very effective with powerful components. However, this method is not very accurate in case of a small sample.[15]

Factor Loadings:
Component or factor coefficient are the factor loadings. They are also known as component loadings in PCA. These factor loadings are regarded as correlation coefficient among the factors and variables. This is similar to Pearson's r. Where the component loading is the variance percent which is the variable that is shown by the factor.[16]

Table (4): Component Matrix

	Component					
	1	2	3	4	5	6
Q1	.659			.366		
Q2	.505	.375	-.336			
Q3	.382					
Q4	-.383	-.334		.406		
Q5	.418	.599				
Q6			.343	.450		.382
Q7				.301		.684
Q8	.829	-.422				
Q9	.760	-.429				
Q10	.784	-.465				
Q11		.339			.456	-.410
Q12	.367		-.367	.507		
Q13			-.587	.353		
Q14			.605		-.435	
Q15	.498	.545				
Q16				.447	-.530	
Q17			.398			
Q18				.421	.464	-.351

This table (4) represents the component matrix which explains the relationship between each variable, and it is factor. Since the first variable is related to the first factor by 0.659 and it is related to the fourth factor by 0.366. However, the second variable is related to first factor by 0.505 and it is related to the second factor by 0.375 and it is related to the third factor by 0.336, and so on.

The variable communality is the sum of its squared factor loadings. reflects the part of the variance in the variable which is accounted for by the common extracted factors. Therefore, in case the communality of variable is great, the extracted factors should be accounted for a higher rate of the variable's variance. Thus, this shows that the particular variables are represented well through the extracted factors.[17]

Table (5): Communalities

	Initial	Extraction
Q1	1.000	.727
Q2	1.000	.566
Q3	1.000	.387
Q4	1.000	.501
Q5	1.000	.571
Q6	1.000	.558
Q7	1.000	.598
Q8	1.000	.877
Q9	1.000	.796
Q10	1.000	.852
Q11	1.000	.599
Q12	1.000	.566
Q13	1.000	.513
Q14	1.000	.598
Q15	1.000	.554
Q16	1.000	.629
Q17	1.000	.381
Q18	1.000	.575

As is apparent from the table (5), the proportion of variance in each variable represent by the six factors is not the same.
It is observed that table (4) which contains the communalities of variables on the six factors. It is found that the first factor contains most of the higher communalities; that means the most variables exist in the first factor. Frequently, this result is accepted to show that most variable depends on only one factor that explains the variance among variables.
It is also observed that some variable is communalited by two or more factors with similar percentages. This leads to the difficulty of explaining the factors. In this case, we should use the rotation method.

2.3 Rotation method (The final solution):

Soon after factor analysis, rotation was emerged and developed in order to support researchers to show the findings of a FA. Rotation method has several options. Each option depends on several algorithms or ways to accomplish a similar broad aim -simplification of factor structure.There are two main classes of rotation methods; oblique and orthogonal this will be studied in this paper, which include the angle found between Y and X axes. Uncorrelated factors are produced by orthogonal rotations in which a 90° angle is maintained between axes. For the oblique method, they permit the factors to be correlation together; that mean the permission of Y and X axes to have a different angle other than 90°.[18]

Orthogonal rotation has Many limitation and advantages when rotation strategy is employed. First, some factors keep unconnected with each other completely and simpler to be analyzed. Second, factor structure and factor pattern matrix are equal. Therefore, a single matrix of association has to be estimated.[19]

One of the most popular orthogonal rotations is varimax rotation. The varimax rotation has an advantage. It is to make the variances of the loading that exist in the factors much greater. However, the high differences between low and high loading on a certain factor. Lower loadings become lower while higher loadings become higher. As a result, varimax rotation becomes good to achieve simple structure. However, it is not suitable for deciding the overall factor as it distributes the variance of major factors on the lesser factor.[20]

Table (6): Rotated Component Matrix

	Component					
	1	2	3	4	5	6
Q1		.655		.404		
Q2		.662				
Q3	.497					
Q4		-.479		.467		
Q5		.716				
Q6						.682

Q7						.738
Q8	.919					
Q9	.879					
Q10	.915					
Q11					.747	
Q12		.359	.535	.320		
Q13			.708			
Q14			-.632	.404		
Q15		.721				
Q16				.764		
Q17		.311				.396
Q18					.738	
Extraction Method: Principal Component Analysis. Rotation Method: Varimax with Kaiser Normalization.[a]						
a. Rotation converged in 11 iterations.						

By comparing table (4) with table (6), it is observed that the rotation that use varimax method redistributes the variance which each factor explains. This means that it changes communalities with it is high and low factor. Consequently, the rotation redistributes the variables on the six factors. As it is observed, the age variable is deleted from the first factor and it transported to the second factor, and it increases the loading on the fourth factor.

Each component is a linear combination of degrees on the original variables as follows:

$$\boldsymbol{F_i = b_{i1}Q_1 + b_{i2}Q_2 + \cdots + b_{ij}Q_j}$$

Where
i: Factor
b_i : Factor coefficients
j: number of variable

$$F_1 = 0.497Q_3 + 0.919Q_8 + 0.879Q_9 + 0.915Q_{10}$$

This reflects that this factor is the most communality with Q_{10}

$$F_2 = 0.655Q_1 + 0.662Q_2 - 0.479Q_4 + 0.716Q_5 + 0.359Q_{12} + 0.721Q_{15} + 0.311Q_{17}$$

This reflects that this factor is the most communality with $Q_1 and\ Q_5$

$$F_3 = 0.535Q_{12} + 0.708Q_{13} - 0.632Q_{14}$$

This reflects that this factor is the most communality with Q_{13}

$$F_4 = 0.404Q_1 + 0.467Q_4 + 0.320Q_{12} + 0.404Q_{14} + 0.764Q_{16}$$

This reflects that this factor is the most communality with Q_{16}

$$F_5 = 0.747Q_{11} + 0.783Q_{18}$$

This reflects that this factor is the most communality with Q_{18}

$$F_6 = 0.682Q_6 + 0.738Q_7 + 0.396Q_{17}$$

This reflects that this factor is the most communality with Q_7

III. Conclusion and Recommendation

In this paper, 18 variables decreased to 6 factors which facilitates finding the appropriate solutions to lower violence or to limit it. The first factor of these six factors is the work factor and the eigenvalue is (3.528). This factor explains the percentage (19.600%) from the total variance. The work factor is the most important in causing violence. The second factor is the family factor who's the eigenvalue is (1.973) and it explains the percentage (10.961%) from the total variance. The third factor is the frequency of committing violence which has an eigenvalue is (1.533) and it explains the percentage (8.516%) from the total variance. The fourth factor is the abuser who has the eigenvalue is (1.453) and it explains the percentage (8.075%) from the total variance. The fifth factor is the income which has an eigenvalue is (1.211) and it explains the percentage (6.726%) from the total variance. Finally, the surroundings that has an eigenvalue is (1.150) and it explains the percentage (6.392%) from the total variance.There are other methods that can be used for extraction and rotation and contrasting the results with the use methods in this study.

After getting these results, we recommend activating certain rules for women protection and to make the society aware of women's rights through the social media, educational institutes including school and universities and to operates the committees for protecting women.

References

[1]. Russo, N. F., &Pirlott, A. (2006). Gender- Based Violence: Concepts, Methods, and Findings. *Annals of the new york academy of sciences*, *1087*(1), 178-205.

[2]. Van der Straten, A., King, R., Grinstead, O., Vittinghoff, E., Serufilira, A., & Allen, S. (1998). Sexual coercion, physical violence, and HIV infection among women in steady relationships in Kigali, Rwanda. *AIDS and Behavior*, *2*(1), 61-73.

[3]. Williams, B., Onsman, A., & Brown, T. (2010). Exploratory factor analysis: A five-step guide for novices. *Australasian Journal of Paramedicine*, *8*(3).

[4]. Henson, R. K., & Roberts, J. K. (2006). Use of exploratory factor analysis in published research: Common errors and some comment on improved practice. *Educational and Psychological measurement*, *66*(3), 393-416.

[5]. Watkins, M. W. (2018). Exploratory factor analysis: A guide to best practice. *Journal of Black Psychology*, *44*(3), 219-246.

[6]. Field, A. (2000). *Discovering Statistics using SPSS for Windows*. London – Thousand Oaks –New Delhi: Sage publications.

[7]. Hair Jr, J. F., Wolfinbarger, M., Money, A. H., Samouel, P., & Page, M. J. (2015). *Essentials of business research methods*. Routledge.

[8]. Arsham, H., &Lovric, M. (2011). Bartlett's test. *International encyclopedia of statistical science*, 87-88.

[9]. Stevens, J. P. (2002). *Applied multivariate statistics for the social sciences* (4th ed.). Mahwah, NJ: Lawrence Erlbaum.

[10]. Ngure, J. N., Kihoro, J. M., &Waititu, A. (2015). Principal component and principal axis factoring of factors associated with high population in urban areas: a case study of Juja and Thika, Kenya. *American Journal of Theoretical and Applied Statistics*, *4*(4), 258-263.

[11]. Hutcheson, G. D., &Sofroniou, N. (1999). *The multivariate social scientist: Introductory statistics using generalized linear models*. Sage.

[12]. Boslaugh, S. (Ed.). (2007). *Encyclopedia of epidemiology*. Sage Publications.

[13]. Tabachnick , B. G., &Fidell, L. S. (2001). *Using multivariate statistics* (4th ed.) Needham Heights., MA: Pearson.

[14]. Velicer, W. F., Eaton, C. A., & Fava, J. L. (2000). Construct explication through factor or component analysis: A review and evaluation of alternative procedures for determining the number of factors or components. In *Problems and solutions in human assessment* (pp. 41-71). Springer, Boston, MA.

[15]. Zwick, R. W., &Velicer, W. F. (1982). Factors influencing four rules of determining the number of components to retain. *Multivariate Behavioral Research, 17*,253-269.

[16]. Ivancevic, T. T., Jovanovic, B., Djukic, S., Djukic, M., & Markovic, S. (2008). *Complex sports biodynamics: With practical applications in tennis* (Vol. 2). Springer Science & Business Media.

[17]. Camara, A. H. (2014). The importance of factor analysis in quantitative and qualitative X-ray diffraction phase analysis. *Koroze a ochranamateriálu*, *58*(2), 52-58

[18]. Osborne, J. W. (2015). What is rotating in exploratory factor analysis. *Practical assessment, research & evaluation*, *20*(2), 1-7.

[19]. Kieffer, K. M. (1998). Orthogonal versus Oblique Factor Rotation: A Review of the Literature regarding the Pros and Cons.

[20]. Costello, A. B., & Osborne, J. W. (2005). Best practices in exploratory factor analysis: four recommendations for getting the most from your analysis. Pract Assess Res Eval 2005; 10. *URL http://pareonline. net/getvn. asp*, *10*(7).

Hadeel F. Alrowithy. " Exploratory Factor analysis: Domestic Violence on Women In Saudi Arabia." IOSR Journal of Mathematics (IOSR-JM) 15.5 (2019): 33-40.