Anatolian Journal of Education
e-ISSN: 2547-9652
April 2019 ● Vol.4, No.1
www.e-aje.net
pp. 59-76

Exploratory Factor Analysis for TPACK among Mathematics Teachers: Why, What and How

Fereshteh Zeynivandnezhad
Organization for Educational Research and Planning, *fzeynivand@gmail.com*

Fatemeh Rashed
Azad University (Neyshabor Branch), Neyshabor, Iran, *rashed_um@yahoo.com*

Arman Kanooni
The Johns Hopkins University (Advanced Academic Programs), *kanooni@sprynet.com*

Factor analysis is a statistical technique that is widely used in psychology and social sciences. Using computers and statistical packages, implementation of multivariate factor analysis and other multivariate methods becomes possible for researchers. Exploratory factor analysis and confirmatory factor analysis are applied in different studies; however, exploratory factor analysis as one of the most important data analysis methods is not well understood. Despite the rising number of researchers who understand it many researchers in various fields still have a false view. This paper reviews methods based on reliable sources and provides a practical guide to conduct an exploratory factor analysis. This paper presents step-by-step of different stages of exploratory factor analysis using SPSS and Montecarlo software. This study could pave the way for students, researchers and teachers who want to use heuristic analysis in their studies of different types of questions, purpose, method of analysis and report writing.

Keywords: factor analysis, exploratory factor analysis, confirmatory factor analysis, principal component analysis

INTRODUCTION

Factor analysis (FA) is a broad term that includes a range of statistical techniques that make it possible to estimate about the total population. This estimate is achieved by a variety of observed variables and relationships between them (Akyuz, 2018; Gorsuch, 1983; Kline, 2011; Matsunaga, 2010). In other words, the goal of factor analysis is to summarize the relationships between variables which enable to conceptualize the phenomena under study (Hair *et al*., 2010; Gorsuch, 1983). In the strict sense of the word, factor analysis responds to research questions of validity as stated that the heart of the management of psychological constructs is factor analysis (Nunnally, 1978). Psychological constructs are variables such as love, motivation, happiness and satisfaction that are not directly measurable (Pallant, 2011; Kline, 2011). In other words, factor analysis is a diagnostic tool for the evaluation of data collected in line with the theoretical model, or the expected target constructs under study. In this case, were the measures used in fact able to measure what they claim? Factor analysis is of great importance in the social sciences and interdisciplinary studies (Yip and Tse, 2019); however, one of the most important methods of statistical analysis is not understood and is not used correctly (Matsunaga, 2010).

Although the number of researchers that understood the issue is on the rise, the evidence suggests that the overwhelming of researches in various fields still have the wrong perspective (Fabrigar, Wegener,

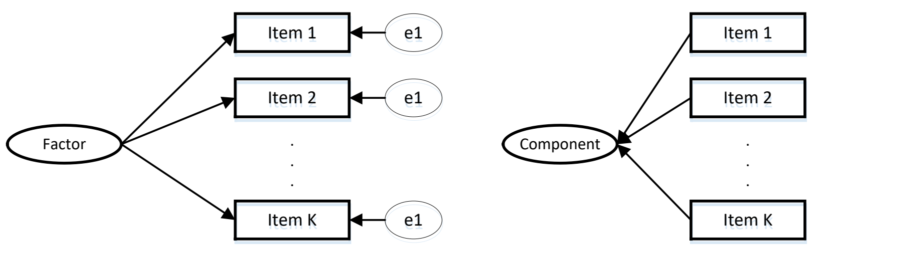
Citation: Zeynivandnezhad, F., Rashed, F., & Kaooni, A. (2019). Exploratory Factor Analysis for TPACK among Mathematics Teachers: Why, What and How. *Anatolian Journal of Education*, 4(1), 59-76. https://doi.org/10.29333/aje.2019.416a

MacCallum, & Strahan, 1999; Henson & Roberts, 2006; Matsunaga, 2010; Park, Dailey & Lemus, 2002; Preacher & MacCallum, 2002). Hoping to make a change in factor analysis, the present article is a practical guide as the best method for performing factor analysis that allow widespread use of different software. There are two ways to do factor analysis, i.e., exploratory factor analysis and confirmatory factor analysis (Hair *et al.*, 2010; Thomason, 2004). Even though both methods are used to test the hidden or latent factors in the data, they play different roles according to the purpose of a research; exploratory factor analysis is to construct the theory and confirmatory factor analysis is to test the theory (Kline, 2011).

This article is depicting a hybrid approach to exploratory factor analysis. Exploratory factor analysis can be used when researchers have a few ideas regarding the mechanism of the phenomenon under study. Therefore, they lack knowledge about the relationship between the variables. For this reason, the researchers resort to exploratory factor analysis to uncover a set of latent factors (constructs) in order to reconstruct the complexity of the observed data under an essential form. It means that the factor extracted from exploratory factor analysis keeps all the important information of variable of the primary data. For example, the resulting solution regarding the variability of individuals and the covariance between the constructs under study stays constant. In other words, the factor analysis is a tool that helps to create a new theory based on the latent factors which is the best way for a variety of variables and the relationships between them (Henson & Roberts, 2006).

EFA, PCA, CFA

In the principal component analysis (PCA) it is assumed that the observable indicators are evaluated without any error of measurement. The principle component analysis (PCA) and exploratory factor analysis (EFA) are calculated based on the correlation matrix. The correlation matrix is a table that shows the relationship between observed variables both vertically and horizontally. The first (PCA) assumes the value (meaning the full reliability) in the main diagonal matrix elements, while the latter (EFA) considers the reliability estimates. Therefore, PCA is neither theoretically nor statistically is a replacement for EFA. Figure 1 shows the conceptual differences between the FA and PCA. Note that an ellipse shows a latent variable in the population. But each rectangle shows an observed variable in the sample data and each arrow shows a directed path. Note that in the factor analysis the measurement error is assumed, however in the principal component analysis it is not.



Exploratory Factor Analysis – The hidden constructs are measured via the observed indicators

Principal Component Analysis – The observed indicators are reduced to factors.

Figure 1
Differences between EFA and PCA (Matsunaga, 2010)

It should be noted that the factor analysis takes place in two ways the exploratory factor analysis (EFA) and the confirmatory factor analysis. Here the exploratory factor analysis concept is explained in depth. The confirmatory factor analysis (CFA) is used to test an existing theory. First, a basic model and the underlying structure of the target constructs using different methods such as exploratory factor analysis (EFA) are assumed, then the model is assessed whether the data are properly match it (Bandalos, 1996). The relationship between the CFA supposed model and the observable data are calculated in the shadow of proper statistics. For the evaluation of a CFA one uses the values of different indices obtained from the software and the researchers will determine whether this resulting model is close enough to the accepted standards (for example see the articles by Hu & Bentler, 1999, Kline, 2005; Marsh, Hu, & Wen, 2004). Exploratory factor analysis is a superficial assessment of the validity of a model. Therefore, Tabachnick and Fidell (2007) asserted that the exploratory factor analysis helps to reduce the number of indicators to constructs, in other words, the conversion of indicators to constructs, resulting in shorter form of the experimental data. Exploratory factor analysis includes:

The formation of a model where the number of latent variables is discovered rather than the goal becomes to determine the particular theory in the construction of a scale or test. Structure of factors is discovered by modeling each indicator as a function of all common factors rather than the indicator becomes just as a subset of the factors. In this case it is determined which factor has a strong relationship with an indicator, and which factor does not. The pattern of factor loading is where the indicators have the greatest load on a factor and they don't on the others (Pallant, 2011).

Exploratory factor analysis is one of the factor analyses that it is used when a researcher does not have any knowledge of the nature or the number factors (Williams, Brown & Onsman, 2010). As the name implies, the exploratory factor analysis allows the researcher to discover the principle variables to build a theory or model through a set of hidden dimensions via a set of indicators (Thompson, 2004 and Henson & Roberts, 2006). CFA unlike EFA includes the hypothesis based on theory or previous model regarding the number of constructs involved and how those models and constructs provide the best fit.

However, EFA and CFA both try to calculate the variance in a set of observed variables into a smaller set of latent variables, factors or components working, EFA is appropriate to provide a scale and when that there exists little theoretical foundation for specifying the number and pattern of common factors (Hurley *et al.*, 1997; Matsunaga, 2010).

In short, the purposes of the exploratory factor analysis are listed as follows:

- Reduce the number of Variables
- Test structure or relationships between variables
- Detection and assessment of a Unidimensionality theoretical constructs
- evaluating the validity of a scale, test or instrument
- Gain simple analysis and interpretation of the phenomena under study
- Address Multicollinearity (high correlation of two or more variables that are correlated with each other)
- The development of theoretical structures
- Approve or disprove the proposed theories

Figure 2 shows the stages of exploratory factor analysis which includes five steps (Taherdoost, Sahibuddin, & Jalaliyoon, 2014; Williams, Brown, & Onsman, A., 2012)

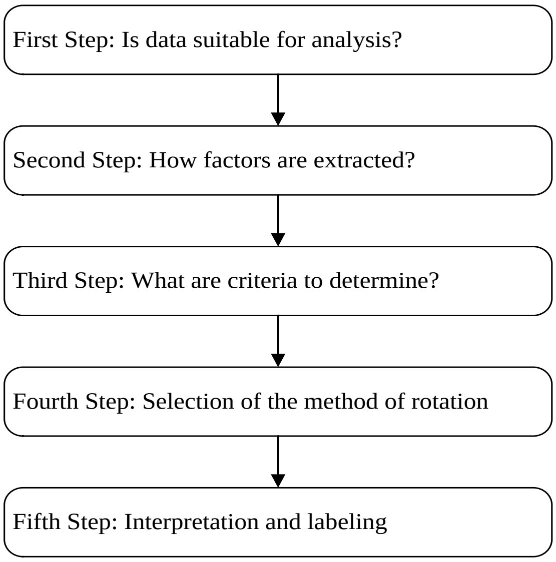


Figure 2
Five Steps to do Exploratory Factor Analysis

Factors of Technological Pedagogical and Content Knowledge (TPACK) using EFA

One supposes in an inquiry to find answer to the following question: what are the factors contributing to TPACK in the sample data? In order to answer this question, a survey questionnaire with eight questions was distributed among 399 teachers enabling to extract the factors related to TPACK based on convenient sampling who were chosen randomly. Exploratory factor analysis was used to determine the number of factors related to TPACK. Three hundred ninety-nine individuals participated in this study where each of the participants received a questionnaire with a specific identification number indicated with the variable Id as part of a column in a SPSS file. The twenty-eight questions of TPACK are indicated by variable Q1 to Q 28. It was used a Likert Scale (Strongly disagree, Disagree, Neutral, Agree and Strongly agree) of five to capture the responses to each questions based on commonly world standard. The table 1 shows the 28 questions about TPACK created based on Alshehri (2012). Data were collected using stratified sampling in 5 provinces among whole of Iran.

Table 1
The Items in the TPACK Questionnaire (Alshehri, 2012)

1 I know how to use different digital technologies.
2 I know how to solve my own technical problems with digital technologies.
3 I frequently play around with digital technologies.
4 I keep up with important new digital technologies.
5 I reason mathematically when I solve problems in my daily life.
6 I can make mathematical connections with the problems outside of mathematics.
7 I am able to communicate mathematically.
8 I use multiple mathematical representations when I solve problems.
9 I know how to adapt lessons to improve student learning.
10 I know how to implement a wide range of instructional approaches.
11 I know how to organize a classroom environment for learning.
12 I know how to assess student performance in a classroom.
13 I have a good understanding of teaching mathematics so that students are able to learn.

14 I have a good understanding of instructional strategies that best represent mathematical topics.
15 I have a good understanding of students' conceptual and practical understanding of mathematical concepts.
16 I have a good understanding of the mathematics curriculum that meets students' needs for learning mathematics.
17 I know how to use digital technologies to represent mathematical ideas.
18 I am able to select certain digital technologies to communicate mathematical processes.
19 I am able to use digital technologies to solve mathematics problems.
20 I am able to use digital technologies to explore mathematical ideas.
21 I am able to identify digital technologies to enhance the teaching approaches for a lesson.
22 I can implement specific digital technologies to support students' learning for a lesson.
23 I think deeply about how digital technologies influence teaching approaches I use in my classroom.
24 I can adapt digital technologies to support learning in my classroom.
25 I know specific topics in mathematics are better learned when taught through an integration of digital technologies with my instructional approaches.
26 I can identify specific topics in the mathematics curriculum where specific digital technologies are helpful in guiding student learning in the classroom.
27 I can use strategies that combine mathematical content, digital technologies and teaching approaches to support students' understandings and thinking as they are learning mathematics.
28 I can select digital technologies to use with specific instructional strategies as I guide students in learning mathematics.

First Step

Are the data suitable for the factor analysis? Before conducting the exploratory factor analysis, it should be made sure that the data are suitable. For this purpose, you should examine the assumptions such as sample size, normal distribution, linearity, outliers and correlation among items. For this purpose, before performing factor analysis one must do the following tasks:

1. Preparing the data
2. The size of the sample
3. Scale of measurement
4. Normality
5. Linearity
6. Outliers
7. Correlation among factors

1. Preparing the data: the outputs by a software package depend on the execution of arranging the data. Therefore, in this regard:

a) The data are controlled, i.e., when the data entry occurs in the software, are there a missing data? And if this case a decision shall be made to deal with it. For this purpose, descriptive statistics such as frequency can be achieved by using the appropriate software.

b) Use the variables that are theoretically in a same group in conceptualization and operationalization stage of developing questionnaire (Pallant, 2011).

2. The size of the sample: Hair, Black, Babin, Anderson, and Tatham (2010) asserted that a researcher shouldn't conduct the factor analysis with a sample of less than 50 observations. They propose that the sample size should be greater than 100 participants. Nunnally (1978) believed that the ratio of a sample to a variable must be ten samples to one variable. But, a general guideline is proposed as follow:

a) Minimum: More than 5 (N) samples for each variable. For example, if there are 20 (p) variables, then the number of samples N should be 100. The ratio of variable to samples is one to five (N:p).

b) Ideal: there are more than 20 samples for each variable, for example if there are 20 variables, and then ideally it would be having 400 samples. So, the ration is 1:20.

c) In general, it is recommended the number of samples greater than 200.

d) Comery and Lee (1992) guidance: they asserted that having 100 samples is weak, while 200 samples are considered fair, 300 samples good, 500 samples very good, 1000 samples and more are excellent.

3. Scale of measurement: All variables for correlation analysis must have ratio scale of measure or the data have at least right interval with Likert scale.

4. Normality: the factor analysis is generally less sensitive to deviations from the normal distribution. If the variables are distributed normally, then it is possible to have better results.

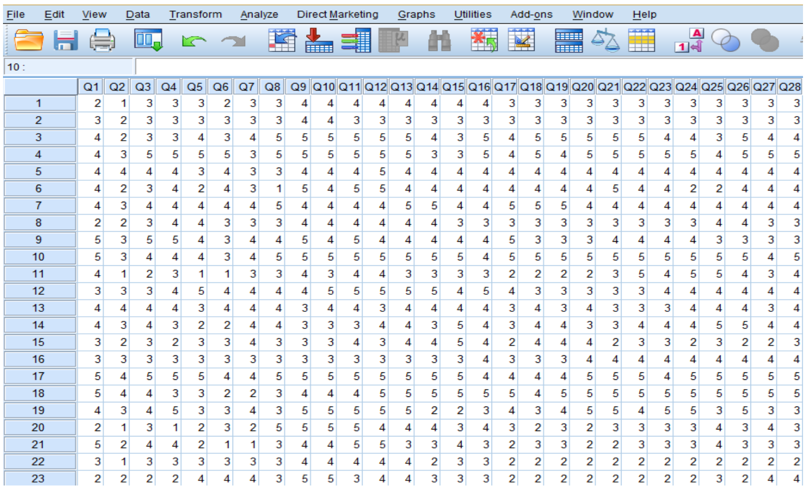
5. Linearity: Since the analysis is based on the correlation between variables, it is important that the linear relationship between the variables to be controlled.

6. Outliers: the factor analysis is sensitive to outliers (Pallant, 2011).

a) The data with bivariate outliers (for this purpose the scatter plot is used).

b) The data with multivariate outliers

It is needed to identify the outliers and then transform or eliminate them. The Figure 3 shows the answers of participants to 28 questions in a spreadsheet within SPSS.



	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9	Q10	Q11	Q12	Q13	Q14	Q15	Q16	Q17	Q18	Q19	Q20	Q21	Q22	Q23	Q24	Q25	Q26	Q27	Q28
1	2	1	3	3	3	2	3	3	4	4	4	4	4	4	4	4	3	3	3	3	3	3	3	3	3	3	3	3
2	3	2	3	3	3	3	3	3	4	4	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3
3	4	2	3	3	4	3	4	5	5	5	5	5	5	4	3	5	4	5	5	5	5	5	4	4	3	5	4	4
4	4	3	5	5	5	5	3	5	5	5	5	5	5	3	3	5	4	5	4	5	5	5	5	5	4	5	5	5
5	4	4	4	4	3	4	3	3	4	4	4	5	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4
6	4	2	3	4	2	4	3	1	5	4	5	5	4	4	4	4	4	4	4	4	5	4	4	2	2	4	4	4
7	4	3	4	4	4	4	4	5	4	4	4	4	5	5	4	4	5	5	5	4	4	4	4	4	4	4	4	4
8	2	2	3	4	4	3	3	3	4	4	4	4	4	4	3	3	3	3	3	3	3	3	3	3	4	4	3	3
9	5	3	5	5	4	3	4	4	5	4	5	4	4	4	4	4	5	3	3	3	4	4	4	4	3	3	3	3
10	5	3	4	4	4	3	4	5	5	5	5	5	5	5	5	4	5	5	5	5	5	5	5	5	5	5	4	5
11	4	1	2	3	1	1	3	3	4	3	4	4	3	3	3	3	2	2	2	2	3	5	4	5	5	4	3	4
12	3	3	3	4	5	4	4	4	4	5	5	5	5	5	4	5	4	3	3	3	3	3	4	4	4	4	4	4
13	4	4	4	4	3	4	4	4	3	4	4	3	4	4	4	4	3	4	3	4	3	3	3	4	4	4	3	4
14	4	3	4	3	2	2	4	4	3	3	3	4	4	3	5	4	3	4	4	3	3	4	4	4	5	5	4	4
15	3	2	3	2	3	3	4	3	3	3	4	3	4	4	5	4	2	4	4	4	2	3	3	2	3	2	2	3
16	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	4	3	3	3	4	4	4	4	4	4	4	4	4
17	5	4	5	5	5	4	4	5	5	5	5	5	5	5	5	4	4	4	4	5	5	5	4	5	5	5	5	5
18	5	4	4	3	3	2	2	3	4	4	4	5	5	5	5	5	5	4	5	5	5	5	5	5	5	5	5	5
19	4	3	4	5	3	3	4	3	5	5	5	5	5	2	2	3	4	3	4	5	5	4	5	5	3	5	3	3
20	2	1	3	1	2	3	2	5	5	5	5	4	4	4	3	4	3	2	3	2	3	3	3	3	4	3	4	3
21	5	2	4	4	2	1	1	3	4	4	5	5	3	3	4	3	2	3	3	2	2	3	3	3	4	3	3	3
22	3	1	3	3	3	3	3	3	4	4	4	4	4	2	3	3	2	2	2	2	2	2	2	2	2	2	2	2
23	2	2	2	2	4	4	4	3	5	5	3	4	4	3	3	3	2	2	2	2	2	2	2	2	3	2	4	4

Figure 3
Data entered in SPSS

SPSS version 23 was utilized to examine the normality of the data providing its value for Kurtosis and Skewness. To do so, it was selected the Analyze/Descriptive Statistics/Explore to obtain the values for Kurtosis and Skewness and the output showed that the Kurtosis is less than 7 and the skewness is less than 2. Based on Kline (2011) the closer the value of Kurtosis is to zero better it is and the quantity of Skewness between -1 and -1 indicates that there is not much deviation from the normal curve as shown on the table 2. However, the researchers also can decide on the removal or maintenance of the indicators.

Table 2
Verification of the Normality of the Data

	N	Min	Max	Mean	Std. Deviation	Skewness		Kurtosis	
	Statistic	Statistic	Statistic	Statistic	Statistic	Statistic	Std. Error	Statistic	Std. Error
Q1	399	1	5	3.55	1.168	-.632	.122	-.303	.244
Q2	399	1	5	2.92	1.180	.007	.122	-.740	.244
Q3	399	1	5	3.20	1.217	-.264	.122	-.770	.244
Q4	399	1	5	3.44	1.167	-.405	.122	-.602	.244
Q5	399	1	5	3.48	1.091	-.347	.122	-.496	.244
Q6	399	1	5	3.27	1.111	-.256	.122	-.610	.244
Q7	399	1	5	3.35	1.140	-.301	.122	-.660	.244
Q8	399	1	5	3.77	1.110	-.751	.122	-.026	.244
Q9	399	1	5	3.90	.931	-.888	.122	.967	.244
Q10	399	1	5	3.86	.991	-.811	.122	.385	.244
Q11	399	1	5	3.93	.935	-.940	.122	1.041	.244
Q12	399	1	5	3.83	.954	-.782	.122	.499	.244
Q13	399	1	5	3.85	.978	-.837	.122	.496	.244
Q14	399	1	5	3.66	.990	-.552	.122	.017	.244
Q15	399	1	5	3.71	.987	-.672	.122	.328	.244
Q16	399	1	5	3.64	.995	-.621	.122	.185	.244
Q17	399	1	5	3.29	1.081	-.355	.122	-.335	.244
Q18	399	1	5	3.27	1.109	-.316	.122	-.481	.244
Q19	399	1	5	3.26	1.146	-.284	.122	-.599	.244
Q20	399	1	5	3.08	1.176	-.071	.122	-.710	.244
Q21	399	1	5	3.20	1.136	-.283	.122	-.565	.244
Q22	399	1	5	3.29	1.115	-.386	.122	-.502	.244
Q23	399	1	5	3.33	1.166	-.436	.122	-.574	.244
Q24	399	1	5	3.49	1.114	-.521	.122	-.359	.244
Q25	399	1	5	3.67	1.106	-.593	.122	-.274	.244
Q26	399	1	5	3.44	1.142	-.348	.122	-.591	.244
Q27	399	1	5	3.32	1.152	-.385	.122	-.524	.244
Q28	399	1	5	3.36	1.106	-.381	.122	-.401	.244
N	399								

To test the linearity, via Graphs/Scatter/Dot checking the linear relationship of item to each other are examined (Figure 4).

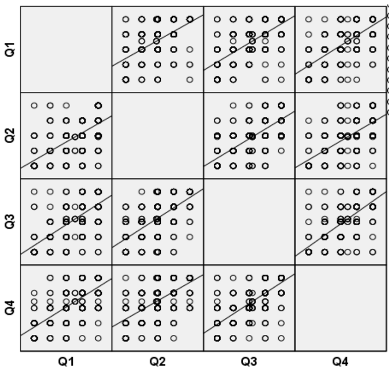


Figure 4
Test of Linearity of Variables with Each Other

As shown in the figure 4, the charts don't exhibit significant changes. Hence, the relationships between the questions in TPACK questionnaire are linear to each other, now; it can be done the exploratory factor analysis. The step two to step five can be done using SPSS.

7. Correlation among items (Factorability): The Correlation Matrix ought to be used in process of showing relationships between variables (Williams, Brown, & Onsman, 2010). Henson and Roberts (2006) indicate that the correlation matrix is one of the most important methods for factor analysis among the researchers. Tabachnic and Fidell (2007) recommended the correlation matrix (shown with the symbol of r) which should have a value of more than 0.3. Hair *et. al* (2010) considered the power of load factor equal to 0.3 as low, 0.4 as important and 0.5 as significant. If there isn't any value greater than 0.3, the researcher must think that is the statistical method of factor analysis suitable for use in his or her research? In other word the correlation value of 0.3 indicates that the factors have approximately 30% relationship to the data or one third of variables have a shared variance therefore this is impractical to determine whether are the variables correlated with each other or with other variables or not? For this purpose, the following criteria are used to determine the contributing factors. Are there some correlation values greater than 0.3? If this is the case then conduct the factor analysis. The table 3 shows an example of a correlation matrix having the correlation values of more than 0.3.

Table 3
Correlation Matrix

		Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9
Correlation	Q1	1.000	.603	.661	.665	.204	.188	.261	.361	.317
	Q2	.603	1.000	.663	.616	.241	.266	.289	.286	.208
	Q3	.661	.663	1.000	.722	.251	.267	.298	.348	.271
	Q4	.665	.616	.722	1.000	.351	.324	.317	.405	.347
	Q5	.204	.241	.251	.351	1.000	.730	.587	.479	.403
	Q6	.188	.266	.267	.324	.730	1.000	.686	.469	.364
	Q7	.261	.289	.298	.317	.587	.686	1.000	.522	.358
	Q8	.361	.286	.348	.405	.479	.469	.522	1.000	.532
	Q9	.317	.208	.271	.347	.403	.364	.358	.532	1.000

The test of Anti-Image Matrices: The values of diagonal of Anti-Image matrix are more than 0.5 as shown in the table 4. The variables with the correlation values of less than 0.5 are excluded from the analysis since they have low correlation with other variables.

Table 4
Anti-Image Matrices

Anti-image Correlation	Q1	.964a	-.177	-.183	-.189	.050	.086	-.040	-.004
	Q2	-.177	.965a	-.250	-.099	-.008	-.041	-.069	.051
	Q3	-.183	-.250	.960a	-.322	.011	.004	-.042	-.042
	Q4	-.189	-.099	-.322	.965a	-.108	-.069	.069	-.026
	Q5	.050	-.008	.011	-.108	.910a	-.516	-.056	-.064
	Q6	.086	-.041	.004	-.069	-.516	.855a	-.433	-.049
	Q7	-.040	-.069	-.042	.069	-.056	-.433	.910a	-.216
	Q8	-.004	.051	-.042	-.026	-.064	-.049	-.216	.960a
	Q9	-.027	.031	-.044	-.012	-.009	-.078	.106	-.202
	Q10	-.043	-.057	.055	-.069	-.015	.105	-.066	-.066
	Q11	.016	.011	.072	-.072	-.036	.071	-.037	-.011
	Q12	-.067	.032	-.046	-.001	-.011	.018	-.043	-.019
	Q13	-.044	.061	.027	.078	-.122	.056	-.083	-.078
	Q14	.109	-.106	.012	-.011	.079	-.167	.084	.039
	Q15	-.131	-.013	.109	.000	-.113	.091	-.150	-.117
	Q16	.061	.107	-.045	.033	.006	-.117	.049	.111

a. Measures of Sampling Adequacy(MSA)

Measures of Sampling Adequacy: This assesses the overall indicators of correlation factors which include the measurement of Bartlett's Test of Sphericity (1954), and the significant and/or Kaiser-Mayer Olkin (KMO) Measure of Sampling Adequacy > .5 or .6. If the measurement of Bartlett is less than 0.05 then this is significant. For example, the table 5 shows such results.

Table 5
KMO and Bartlett's Test

KMO and Bartlett's Test		
Kaiser-Meyer-Olkin Measure of Sampling Adequacy.		.954
Bartlett's Test of Sphericity	Approx. Chi-Square	9054.295
	df	378
	Sig.	.000

In summary, the adequacy of sampling includes some of the following measures that will help to determine the correlation between variables:

- A number of correlation values are larger than 0.3.
- The amounts on the diagonal are more than 0.5.
- The Bartlett' Test produces a significant value.
- The value for KMO measure is larger than 0.5 or 0.6.

Second step: How factors are extracted?

Principal component analysis (No theory has a priority or there is no preexisting model). Pett, Lackey, & Sullivan (2003) proposed this method as one of the first solution to this problem.

- Principal axis factoring
- Maximum likelihood
- Unweighted least squares
- Generalized least squares
- Alpha factoring
- Image factoring

Third step: What criteria are helping to determine the extraction of factors?

The reduction of the elevated number of indicators to factors- Multiple criteria – Factorability (Hair et al., 2010), Kaiser criteria (Eigenvalue greater than 1 and Kaiser, 1960), Scree plot test (Cattell, 1966), the cumulative percentage of variance extracted and parallel analysis (Horn, 1965)

The cumulative percentage of variance and eigenvalue greater than 1

Based on Hair *et al.* (2010), in the natural sciences once the factors are obtained, they can explain at least 95 percent of the variance. But in the human sciences, usually about 50 to 60 percent of variance can be explained. The communalities show the usefulness of describing the extracted components based on the degree of variance of the measured variable (Thompson, 2004). This communality shows the R square among the constructs and the indicators. In other word, the communality is equivalent to R^2 in the regression equation. All the indicators should have a load factor of more than 0.5. In the communalities matrix, the initial communalities in a principal component analysis is always equal to 1 since the researcher tries to explain all of the variances in each item (Gable & Wolf. 1993).

The table of Total Variance Explained shows all the real components that are extracted. The components at the beginning are discovered by SPSS version 23 considering the variables that are in the data (Pallant, 2011). The column "% of Variance" in the table of Total Variance Explained states that the variability of each of the components. In order to determine the number of variables, the eigenvalue larger than 1 are considered (Table 6).

Table 6
Total Variance Explained

Component	Initial Eigenvalues			Extraction Sums of Squared Loadings			Rotation Sums of Squared Loadings		
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	13.313	47.545	47.545	13.313	47.545	47.545	9.570	34.178	34.178
2	3.631	12.968	60.513	3.631	12.968	60.513	5.134	18.336	52.514
3	1.466	5.237	65.749	1.466	5.237	65.749	3.126	11.165	63.678
4	1.123	4.012	69.762	1.123	4.012	69.762	1.703	6.083	69.762

Extraction Method: Principal Component Analysis.

Scree test

It is passed a straight line from the eigenvalues in the Scree Plot and it takes into account where this line is mutated. This point is where the curve is broken. The numbers above the broken place show the number of obtained factors. If the Scree Plot is very busy, then it becomes difficult to interpret the number of factors and hence it needs to resort to other operations to extract the factors.

Parallel Analysis

In the parallel analysis the eigenvalues obtained from the principal component analysis are distorted randomly and in the new matrix the eigenvalues are recalculated again (Table 9) and if the values obtained are less than the previous values, then it is accepted, otherwise it is rejected. This method should be done carefully and for this reason many researchers void all together.

Forth step: Selection of the method of rotation

Orthogonal Varimax (this produces the factor structures that constructs are not correlated to each other).

Oblique rotation (the constructs are correlated to each other), when the data don't have any pre hypothetical relationship, the results are more accurate for research in behavioral science.

Fifth step: Interpreting and labeling

Researcher examines what variables (indicators) are assigned to a variable and based on the variables (indicators) those factors are named. Traditionally, at least there are two or three variables (indicators) on a factor that are loaded that make the interpretation of that factor meaningful. Labeling the factors is an inductive, posteriori, and theoretical process. Providing meaning to the latent factors in the final analysis depends on the definition given by a researcher (Henson & Roberts, 2006). The goal of adding label to the factors is to develop a theoretical and conceptual target (Figure 5).

Steps 2 to 5: it was used the software SPSS version 23 to conduct the steps 2 through 5. The goal was to find the factors contributing to TPACK. As shown in the Figure 5: To start the factor analysis, the Analysis menu and choice the Dimension Reduction and the Factor are selected.

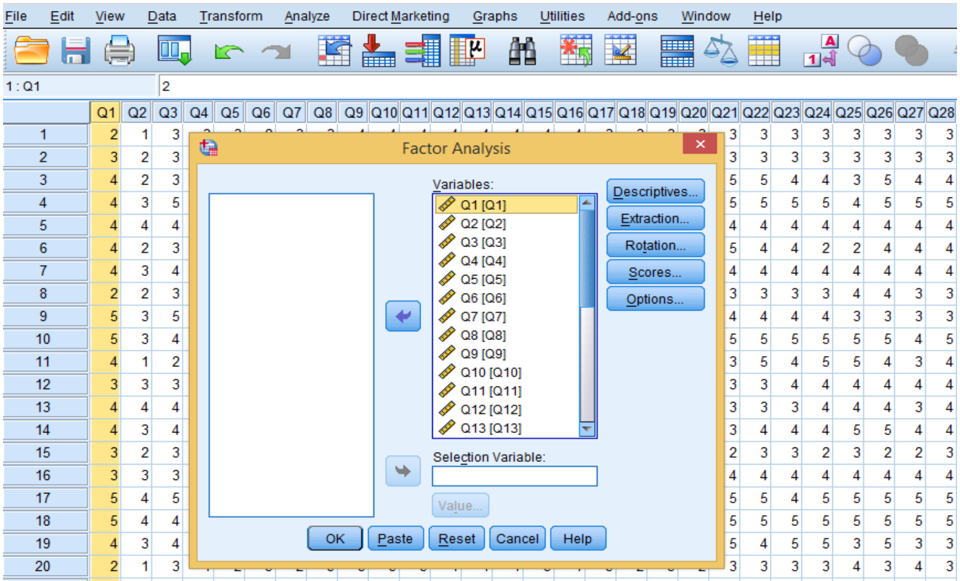


Figure 5
Factor Analysis Using SPSS Version 23

All the five stages of this process are explained in the step 6 are run using SPPS. In these stages it is noted the correlation factors, extraction, and type of rotations. The aforesaid above are easily obtained by using SPSS software which should be paid particular attention to interpret the outputs. There are five icons in the Factor Analysis as shown in the figure 6. These are Description, Extraction, Rotation, Scores, and Options.

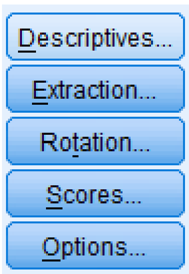


Figure 6
Different Stages of Factor Analysis in SPSS

As mentioned in the five stages of factor analysis, the items related to the rotation and the relationship between factors are conducted to extract the factors. These items can be easily done via the software. In the Descriptive section, the checked box for correlation, anti-image, and KMP are ticked and press continue.

In the Extraction section the principle component is selected as the method of extraction. It must be noted that it is selected this method when there is not any pre supposition about the factors.

Also, it is selected the Correlation Matrix and Scree Plot options and the Extraction section and defined the eigenvalues of more than one. If there is the knowledge about the number of factors based on the literature review, then it is specified in the section of Fixed Number of Factors the number of

Factors to Extract, for example 4 as number of factors. But, generally speaking, there is no knowledge about the number of factors, it is used the special number of 1.

The most used common orthogonal approach is Varimax that strives to reduce the number of variables that have high loadings. On the other hand, the most used common oblique approach is Oblimin Direct. It can be referred to Tabachnick, Fidell, & Osterlind (2001) text book to compare the characteristics of each one. In this article based on Pallant (2011) first the factor analysis is conducted using Varimax and rotation since it was mentioned that it is easier to interpret the results and then the analysis will be carried out using Oblimin which shows the strength of real correlation between variables and also it shows which method is most appropriate (See the Figure 7).

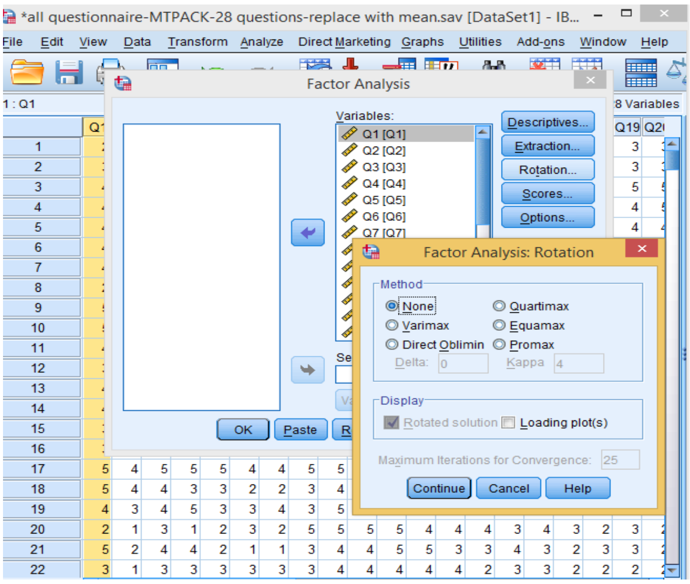


Figure 7

Factor Analysis Rotation Method Based on Varimax Using SPSS

Table 7 shows the extracted indicators loaded for each component. As can be seen, indicators of Q28 to Q2 are loaded on the first component (the load factor of more than 0.5). The Varimax rotation is applied to maximize factor loading of each predictor variables (items) on each component. In general, the principal component analysis with Varimax rotation is used to rank a series of components.

Table 7
Rotated Component Matrix

	Component			
	1	2	3	4
Q28	.845			
Q24	.830			
Q19	.817			
Q21	.813			
Q20	.811			
Q26	.804			
Q27	.794			
Q22	.791			
Q18	.777			
Q17	.771			
Q23	.736			
Q25	.720			
Q3	.652			.514
Q4	.610			.524
Q1	.598			.524
Q2	.578			.540
Q9		.798		
Q10		.798		
Q11		.793		
Q12		.784		
Q13		.698		
Q14		.637		
Q16		.628		
Q15		.555		
Q8				
Q6			.866	
Q5			.793	
Q7			.770	

Extraction Method: Principal Component Analysis.
Rotation Method: Varimax with Kaiser Normalization.a
a. Rotation converged in 5 iterations.

As mentioned before, the five steps are used for the principal component analysis. But, Matsunaga (2010) proposed the hybrid approach to do the exploratory factor analysis. For this purpose, first the principal component analysis is carried out as explained before, in which the initial measures are reduced to a set of ensemble containing the linear combination of variances (Pallant, 2011) and it is used the principal component analysis with Varimax rotation to extract the constructs of each questionnaire. Therefore, the criteria such as Kaiser-Meyer-Olkin (KMO), Barlett's Sphericity Test, eigenvalues, and factor loadings are applied. It can be referred to table 7 where it is addressing these eigenvalues. Then it is conducted a parallel analysis using the MonteCarlo PCA. So an exploratory factor analysis by principal component analysis is done for each questionnaire.

In order to examine the correlation between the components, the factor analysis using Direct Oblimin method is conducted. A close look at the pattern matrix and the component correlation matrix show the factors and the correlation between them. This means that once again the factor analysis is conducted using the Direct Oblimin method. The load factor is shown in the table 7 and the correlation between factors is showed in table 8.If the values are more than 0.8, then the multicollinearity occurs which should be examined closely and eliminated.

Table 8
Component Correlation Matrix

Component	1	2	3	4
1	.796	.472	.323	.198
2	-.538	.689	.441	-.202
3	-.015	-.548	.837	.000
4	-.277	.048	.026	.959

Extraction Method: Principal Component Analysis.
Rotation Method: Varimax with Kaiser Normalization.

As mentioned above the hybrid approach is used for exploratory factor analysis, hence at this stage, it is used parallel analysis to check the number of components. For this purpose, MonteCarlo.exe file is used which is available on the Internet (Figure 8).

The number of variables is 28 and the number of participants is 399. The output of this software is shown in table 10 using MonteCarlo.exe program. The eigenvalues are compared with the Total Variance Explained and Initial Eigen Values. As shown in table 11, the eigenvalues are placed under the Actual eigenvalue from PCA and under the Criterion value from parallel analysis in order to compare the results. If the value obtained from MonteCarlo.exe program is less than the value obtained from SPSS program, then it is accepted the result.

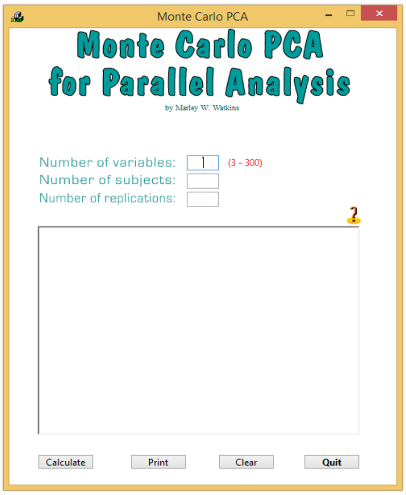


Figure 8
A Screen Shot of MonteCarlo.exe Program

In table 12, there are 3 extracted factors related to items namely TPACK, PK and CK. It worth noting that the factor loading of Q8 was less than .5 therefore it was removed from the analysis.

Table 9
Parallel Analysis

Component Number	Actual eigenvalue from PCA	Random order from parallel analysis	Decision
1	13.313	1.5187	Accept
2	3.631	1.4452	Accept
3	1.466	1.3891	Accept
4	1.123	1.3421	Reject

The report for the exploratory factor analysis is presented as follow: a questionnaire with 28 questions regarding TPACK using the factor analysis with principal components analysis via SPSS was conducted. Before conducting principal component analysis, the suitability of the data for the factor analysis needs to be assessed. The examination of correlation matrix showed a factor of 0.3 and more. The value for the Kaiser-Meyer-Oklin was at 0.954 which is higher than the recommended value of 0.6 (Kaiser, 1970). The examination of Barlett's Test of Sphericity (Barlett, 1954) was significant and supported the correlation coefficient. The principal component analysis revealed three components with the following dispersion 47.545%, 12.968%, and 5.237%. Also, the scree plot of failure curve indicates three principal components. To keep these three components, parallel analysis was carried out. For this purpose, it was utilized the MonteCarlo.exe program producing three eigenvalues which can be compared to eigenvalues generated by SPSS. These values were obtained accidentally via a 28 by 399 matrix. The test showed that the three items on each component are generally strong. The three components explained 63.678% of the dispersion which were 34.178% for the first component, 18.336% for the second, and the 11.165% for the third one. The Figure 9 shows the summary of steps in SPSS.

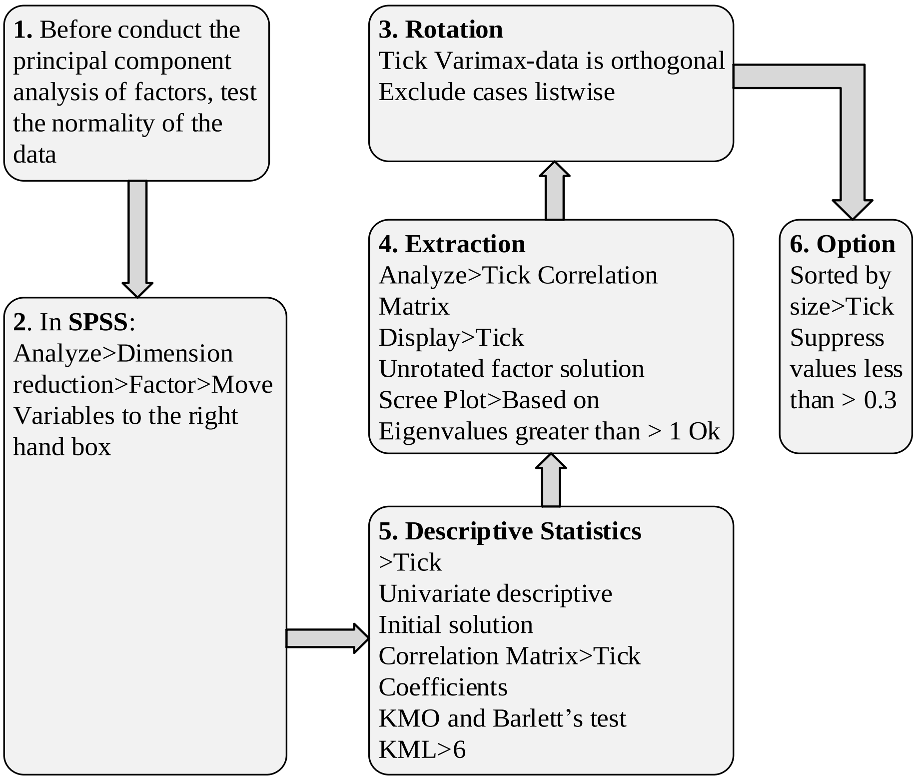


Figure 9
The Process of Component Principal Analysis in SPSS

Once it was obtained the results of the analysis via SPSS to determine the number of factors, the Monte Carlo PCA for parallel analysis is conduct to compare the total variance explained in a table and is decided for the number of factors.

CONCLUSION

This paper outlines the ways to do exploratory factor analysis as a technique for data reduction to factors. First, reviewed the relevant theoretical knowledge and then use the data in an effort to provide practical guidance in this area. The factor analysis is conducted to develop and evaluate the instruments and the measures and also to reduce the large number of relevant variables to the number of possible manageable ones prior to its use in the analysis of structural equations modeling like the other or multiple regression and analysis of multivariate to work. There are two factor analysis approaches including exploratory and confirmatory factor analysis. The first one is to collect and discover the relationships between variables and the second one is about the approval or rejection of a particular theory or assumption about the structure of a set of variables. And despite the fact that the exploratory factor analysis as one of the statistical methods is widely used in psychology research and management, researchers will often be a critical blow in decisions do this analysis. For this purpose, in accordance with the research literature, procedures related to exploratory factor analysis and specific recommendations were presented. Meanwhile, the respective cases by practical step-by-step data were analyzed by SPSS version 23. In this article the concepts such as EFA, PCA and the CFA were expressed with respect to the research literature. Providing relevant concepts can help researchers conduct the exploratory analysis methods for their study design according to the purpose and the appropriate research questions. This article describes recommendations regarding the size of the sample, extraction method, determining the number of factors, rotation method, and method of estimation. It was recommended that a minimum of 20 samples to be there for each variable. However, in any case, the number of samples of more than 200 people is acceptable. Criteria such as correlation coefficient greater than 3.0, Bartlett's test ($P<0.05$) considered significant, and KMO of more than 0.6 were proposed. There are seven methods for extracting the data (Extraction), the most common of which is the PCA. In this method the variables are examined to a set of smaller linear combinations by all variances in variables. However, factor analysis is estimated using a mathematical model, and just shared variance is analyzed. Regarding the decision making about the factors, multiple criteria should be considered. At the beginning the Kaiser's Rule with the eigenvalues of larger than 1 the factors are determined. After that test is resorted to the scree graph that used eigenvalues on the chart connects them according to the order that are acceptable which is higher than this point. The most accurate method for making decisions about the number of factors is the parallel analysis (Horn's PA) that the eigenvalues that are produced accidentally by the software MonteCarlo.exe and the eigenvalues obtained by using SPSS software which are compared together. The values are acceptable when the eigenvalues greater than eigenvalues derived by the application. In practice, this can be done normally after the factors that were analyzed by SPSS were rotated. Rotation is obtained as a mathematical calculation to yield new load factors. Factor rotation has two orthogonal and diagonal approaches. According to the literature review the recommendation is to first using PCA via the Varimax to rotate factors and in the second phase the oblique rotation via Direct Oblimin should be conducted to assess the correlation between the factors. This is a step by step process using the software SPSS executed and should be a practical guide for interdisciplinary researchers. Note that each time one runs a factor analysis an indicator is eliminated with the lowest load factor and again do all the steps of factor analysis or keep the number of factors in the process of factor analysis constant.

REFERENCES

Akyuz, D. (2018). Measuring technological pedagogical content knowledge (TPACK) through performance assessment. *Computers & Education*. *125*, 212-225.

Alshehri, K. A. (2012). The influence of mathematics teachers' knowledge in technology, pedagogy and content (TPACK) on their teaching effectiveness in Saudi public schools. Dissertation Abstracts International: Section A. Humanities and Social Sciences, 74(02), 3541570.

Bandalos, B. (1996). Confirmatory factor analysis. In J. Stevens (Ed.), *Applied multivariate statistics for the social sciences* (3rd ed., pp. 389-420). Mahwah, NJ: LEA.

Bartlett, M.S. (1954) A Note on the Multiplying Factors for Various Chi Square Approximations. *Journal of the Royal Statistical Society, 16*, 296-298.

Cattell, R. B. (1966). The scree test for the number of factors. *Multivariate behavioral research*, *1*(2), 245-276.

Comery, A. L. and H. B. Lee (1992). *A First Course in Factor Analysis*. Hillsdale, NJ, Erlbaum.

Fabrigar, L. R., Wegener, D. T., MacCallum, R. C., & Strahan, E. J. (1999). Evaluating the use of exploratory factor analysis in psychological research. *Psychological methods*, *4*(3), 272.

Gable, R. K., & Wolf, M. B. (1993). *Instrument development in the affective domain: Measuring attitudes and values in corporate and school setting*. Massachusetts: Kluwer Academic Publisher.

Gorsuch, R. L. (1983). Factor analysis (2nd ed.). Hillsdale, NJ: LEA.

Hair, J. F., Black, W. C., Babin, B.J., Anderson, R. E., and Tatham, R. L. (2010). *Multivariate Data Analysis*. New Jersey: Pearson- Prentice Hall.

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*, 393-416.

Horn, J. L. (1965). A rationale and test for the number of factors in factor analysis. *Psychometrika*, *30*(2), 179-185.

Hu, L.-T., & Bentler, P. M. (1999). Cutoff criteria for fit indices in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling, 6*, 1-55.

Hurley, A. E., Scandura, T. A., Schriesheim, C. A., Brannick, M. T., Seers, A., Vandenberg, R. J., & Williams, L. J. (1997). Exploratory and confirmatory factor analysis: Guidelines, issues, and alternatives. *Journal of organizational behavior*, *18*(6), 667-683.

Kaiser, H. F. (1960). The application of electronic computers to factor analysis. *Educational and psychological measurement.*

Kaiser, H., (1970). A second generation Little Jiffy. *Psychometrika*, *35*, 401–15.

Kim, J.-O., & Mueller, C. W. (1978). *Introduction to factor analysis*: *What it is and how to do it.* Newbury Park, CA: Sage.

Kline, R. B. (2005). *Principles and practice of structural equation modeling* (2nd ed.). New York: Guilford

Kline, R. B. (2011). *Principles and practice of structural equation modeling* (3rd ed.). New York: Guilford Press.

Marsh, H. W., Hu, K.-T., & Wen, Z. (2004). In search of Golden Rules: Comment on hypothesis-testing approaches to setting cutoff values for fit indexes and dangers in over generalizing Hu and Bentler's (1999) findings. *Structural Equation Modeling, 11*, 320-341.

Matsunaga, M. (2010). How to Factor-Analyze Your Data Right: Do's, Don'ts, and How-To's. *International Journal of Psychological Research,3*(1), 97-110.

Nunnally, J. C. (1978). *Psychometric theory* (2nd ed.). New York: McGraw-Hill.

Pallant, J. (2011). Multivariate analysis of variance. *SPSS survival manual*. Crows Nest: Allen & Unwin, 20(11), 283-96.

Park, H. S., Dailey, R., & Lemus, D. (2002). The use of exploratory factor analysis and principal components analysis in communication research. *Human Communication Research*, *28*(4), 562-577.

Pett, M., Lackey, N. & Sullivan, J. (2003). *Making sense of factor analysis*. Thousand Oaks: Sage Publications, Inc

Preacher, K. J., & MacCallum, R. C. (2002). Exploratory factor analysis in behavior genetics research: Factor recovery with small sample sizes. *Behavior genetics*, *32*(2), 153-161.

Royston, J. P. (1995). Shapiro-Wilk normality test and P-value. *Applied Statistics*, *44*(4), 547-551.

Tabachnick, B. G., Fidell, L. S., & Osterlind, S. J. (2001). Using multivariate statistics. Boston, MA: Allyn & Bacon.

Tabachnick BG, Fidell LS. (2007), *Using Multivariate Statistics*. Boston: Pearson Education Inc.

Taherdoost, H., Sahibuddin, S., & Jalaliyoon, N. (2014). Exploratory Factor Analysis; Concepts and Theory. In *International Conference on Mathematical- computational and Statistical-Sciences*, Gdansk-Wrzeszcz, Poland.

Thompson, B. (2004). *Exploratory and confirmatory factor analysis*. Washington, DC: American Psychological Association.

Williams, B., Brown, T., & Onsman, A. (2012). Exploratory factor analysis: A five-step guide for novices. *Journal of Emergency Primary Health Care*, *8*(3),1–13.

Yip, T. H. J., & Tse, W. S. (2019). Why hope can reduce negative emotion? Could psychosocial resource be the mediator?. *Psychology, health & medicine*, *24*(2), 193-206.