

HEURISTICS OF USING FACTOR ANALYSIS IN SOCIAL RESEARCH

Kamran Ahmed Siddiqui

University of Dammam, Dammam, Saudi Arabia

Email: KASiddiqui@uod.edu.sa

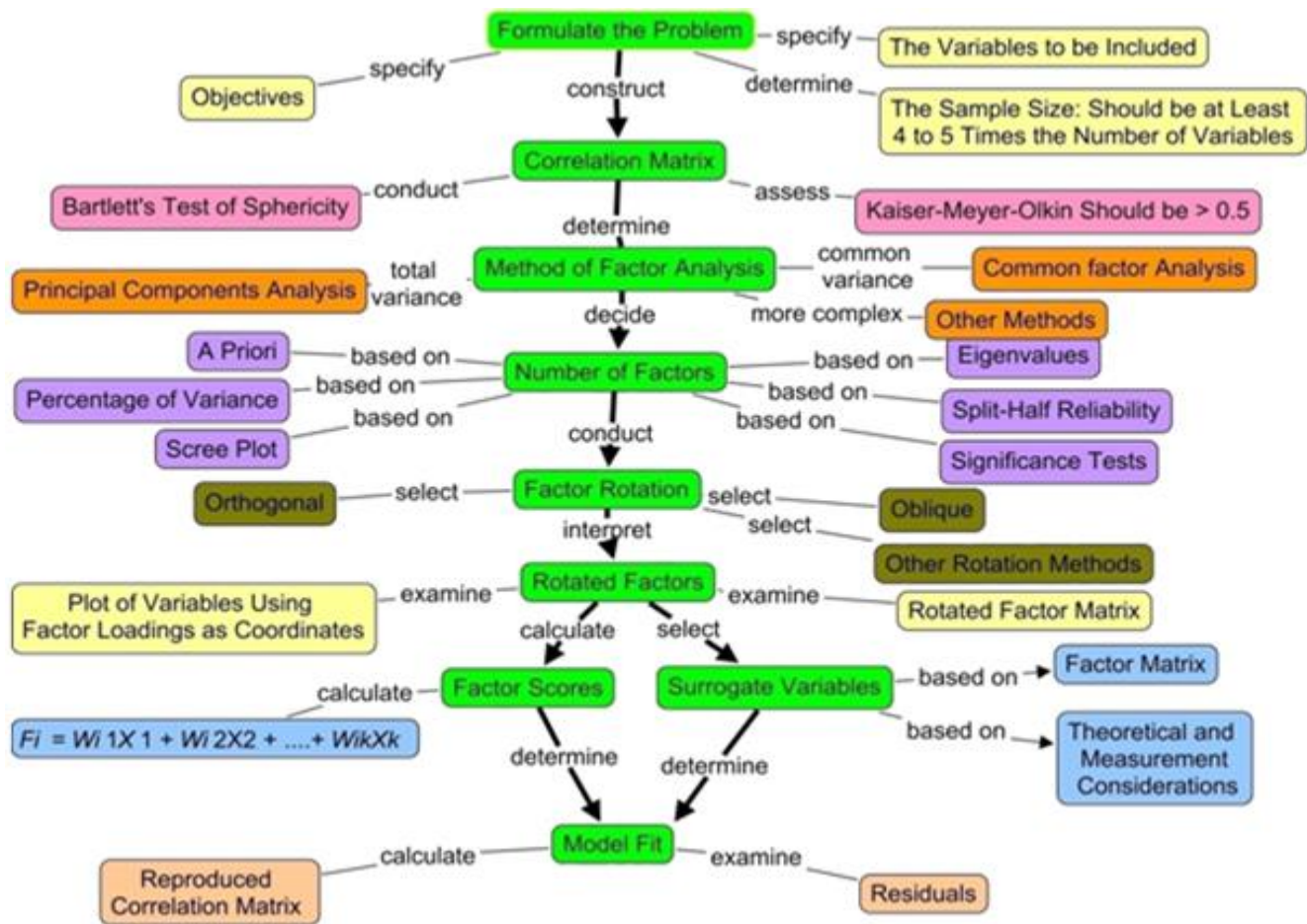
ABSTRACT: This paper aims to provide guidelines for using Factor Analyses in social research. Guidelines are provided for eight mandatory methodological areas; (a) Objectives of using Factor Analysis; (b) Types of Factor Analyses; (c) Factor Order (d) Sample Size requirements (e) Factor Extraction (f) Factor Rotation (g) Factor labelling (h) Summated Score.

I. INTRODUCTION

Employing multivariate statistical techniques in social research projects are always considered as a cumbersome task. Experienced researchers have developed heuristics (rules-of-thumb) for using these techniques in research. This paper offers reliable literature for successful and meaningful application of Factor Analysis in social research in an abridged form.

II. LITERATURE REVIEW

Factor analysis is considered as one of the advanced statistical technique frequently used by psychologists and sociologists [1], and pioneered by renowned psychologists Pearson and Spearman [2]. It is defined as an interdependence multivariate statistical technique used to study the underlying structure and relationships of variables [3]. Concept map for Factor Analysis is shown in Figure 1.



Source: Concept map of Factor Analysis [4]

This article provides guidelines on various methodological areas; (a) Objectives of using Factor Analysis; (b) Types of Factor Analyses; (c) Factor Order; (d) Sample Size requirements (e) Factor Extraction (f) Factor Rotation and (g) Factor labelling.

Objectives: Like other statistical techniques, the starting point in Factor Analysis is to review the research objectives [4]. Objectives of using Factor Analysis are summarized in Table 1. Most often social scientists use EFA to summarize

data [7] or examine the structure/ relationships of variables [8]. Occasionally it has been used to validate the constructs and instruments [9]. In contrast CFA is only used to test the data fitting to proposed model [10]. Validity usage of CFA is also common in social sciences.

Types: Two types of Factor Analyses are summarized in Table 1. Based on the research objectives appropriate category of Factor Analysis needs to be selected.

b) **Factor Order:** Second order factor or higher factors are very common in social sciences and various psychological constructs like intelligence [11], personality [12], advertising [13], and marketing [14] are having higher order factors. A hypothetical structure of Higher Order Factor Analysis is shown in Fig 2 [15].

Table 1 Summary of Types of Factor Analyses and their characteristics

	Exploratory Factor Analysis (EFA)	Confirmatory Factor Analysis (CFA)
Defined[4]	Used to identify complex interrelationships among items and groups that are part of unified concepts.	More complex approach that tests the hypothesis and helps to build models that fits the data..
Objectives	<ul style="list-style-type: none"> Summarizes / reduces the number of variables [5] Examines the structure/relationships between variables [5] Assesses the uni-dimensionality of constructs [5] Evaluates the construct validity of a scale/ instrument [5] Helps to develop theoretical constructs [5] Helps to prove/disprove proposed theories [5] Helps to develop parsimonious interpretation [5] Addresses multi-collinearity [4] 	<ul style="list-style-type: none"> Tests the data to fit a hypothesized model [6] Evaluates the construct validity of a scale/ instrument [6] Provides convergent & discriminant validity [6]
'A priori' knowledge[4]	Can work with and without 'a priori' knowledge /assumptions about factors and their relationships. With 'a priori' knowledge the researcher can solicit exact number of factors needed while without "a priori knowledge" researchers use the EFA procedures to diagnose the number of factors and their relationships.	Can only work with 'a priori' knowledge about factors and their relationships. With 'a priori' knowledge the hypothesized models are tested against actual data, and the analysis would demonstrate factors and their relationships in graphical form.
Software Used[5]	General purpose statistical software such as SPSS, SAS, and Stata.	Specialized software for CFA are used such as AMOS; LISREL, and Mplus

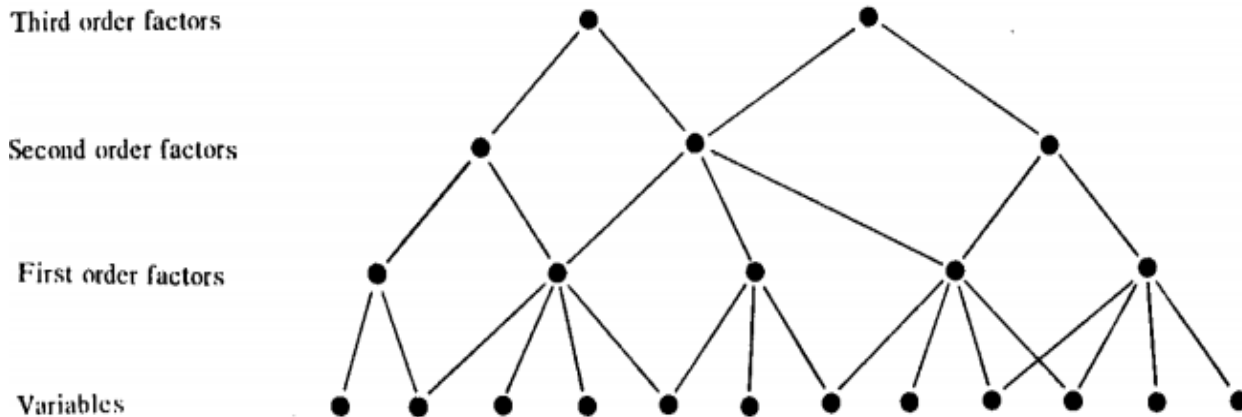


Fig. 1 Hypothetical Structure of Higher Order Factor Analysis

Source: Hypothetical Structure of Higher Order Factor Analysis [7]

c) **Sample Size Requirements:** Factor analysis is sensitive to sample size [16], and even small variations in sample size may yield different results [17]. Different levels of sample size are ranked below [18]:

Table 2 Sample size level and rating

Sample Size	Rating
50	Very Poor
100	Poor
200	Fair
300	Good
500	Very Good
1,000	Excellent

Most of guidelines for Factor Analysis advocate for larger sample size; (a) at least 200 respondents must be sampled in

order to attain a stable solution through factor analysis [19] (b) at least 400 or greater to produce undistorted results [20] (c) appropriate sample sizes depend upon the numbers of items available for factor analysis [21] as shown below:

Table 3 . Appropriate sample size for factor analysis

Sample Size	Number of items
200	10
250	25
400	90
700-1000	500

Suitability of sample size for factor analysis should be tested by Kaiser-Meyer-Olkin Measure of Sampling Adequacy (KMO) [22]. It recommended when the cases to variable ratio are less than 1:5. The KMO ranges from 0 to 1, with 0.50 considered suitable for factor analysis [23].

d) **Factor Extraction:** It involves meeting three groups of criteria. Firstly the Extraction method, secondly factor extraction criteria and thirdly factor loading criteria [24]. Various factor extraction methods are available in different software programs. Normally social scientists work with default methods. Table 4 shows that the default extraction methods. For EFA is the default method is PCA while ML is default method for CFA.

Table 4 Factor Extraction Methods for Social Research

EFA/SPSS - Principal Component Analysis (PCA) [25]
CFA/AMOS - Maximum Likelihood method (ML) [26]

Factor extraction criteria [27] are summarized in table 5 which requires multiple criteria to be used.

Table 5 Factor Extraction Criteria

- The eigenvalue of a factor should be ≥ 1 .
- Scree test should reflect the the appropriate number of factors.
- A priori' knowledge about the theoretical constructs and their relationships.

The factor loading criteria [28] are summarized in table 6 which must be considered extracting factors from data.

Table 6 Factor Loading Criteria

- Each factor must have at least three item loadings ≤ 0.3 ;
- Individual items must have at least one loading ≤ 0.3 ;
- In case of cross loadings the item will be placed only in the factor on which it has higher factor loadings.
- If cross loading is ≥ 0.5 on both factors, the item must be considered for deletion.

e) **Factor Rotation:** There are two basic types of rotation: (a) Orthogonal and (b) Oblique. Orthogonal means the factors are assumed to be uncorrelated with one another. Oblique rotation derives factor loadings based on the assumption that the factors are correlated, and this is probably most likely the case for most of the social science research projects. It has been observed that social science constructs are often correlated. Results often reveal that most items load on the early factors, and usually, of having many items load substantially on more than one factor. It requires Rotation techniques to make the output more 'Simple Structure': For PCA extraction varimax rotation is the most popular rotation method while for ML extraction method Direct Oblimn rotation method is recommended for social sciences.

f) **Factor labelling:** In order to assign some meaning to the factor solution, each factor must be labelled according to the nature of the items loading on each factor in line with standard practice. Social scientist are known for their skills for summarizing the items into fewer factors and then name the factors as true representatives of all or majority of items loaded.

g) **Summated Scores:** In order to assign some meaning to the factor solution, each factor must be labelled according to the nature of the items loading on each factor in line with standard practice. Social scientist are known for their skills for summarizing the items into fewer factors and then name the factors as true representatives of all or majority of items loaded.

III. PRACTICAL IMPLICATIONS

In order to make their research argument a bit more valid and reliable social scientists especially research students are

always inclined to use multivariate statistical techniques in their social research projects. This research paper outlines the basics concepts of using Factor Analysis in social research projects. This paper aims towards young and inexperienced researchers so that they can get familiarize with the basic concepts about the Factor Analysis.

IV. RESEARCH LIMITATIONS

Exposure of author is limited to only social sciences therefore scientists from other disciplines may consult experts in their own fields. Secondly this research paper does not provide any mathematical computations for Factor Analysis rather it only outlines the basic points of using Factor Analysis technique in research projects.

V. ORIGINALITY/VALUE

Although there are other studies providing a numerical solution to the problem; this paper answers to the demanding question from less-experienced researchers who would like to have some rule of thumb decide upon. By contributing to the body of knowledge in this area, this research adds significant value.

REFERENCES

- [1] Pervin, L.A., John, O.P., "Personality – Theory and Research", 7th edition, NY: John Wiley & Sons (1997)
- [2] Gorsuch, R. L., "Factor Analysis", 2nd edition, Hillsdale,NJ: Lawrence Erlbaum Associates (1983).
- [3] Hair, J.F., Black, W.C., Babin, B.J., Anderson, R.E., "Multivariate Data Analysis", 7th Edition, NJ: Prentice Hall, (2010)
- [4] Malhotra, N. K. Marketing research: An applied orientation. Upper Saddle River, NJ: Pearson/Prentice Hall (2007).
- [5] Hair, J.F., Black, W.C., Babin, B.J., Anderson, R.E., "Multivariate Data Analysis", 7th Edition, NJ: Prentice Hall, (2010)
- [6] Williams, B., Brown, T., & Onsman, A. "Exploratory factor analysis: A five-step guide for novices". *Australasian Journal of Paramedicine*, **8**(3). (2010).
- [7] Brown, T.A., Confirmatory Factor Analysis for Applied Research, 2nd Edition, NY: Guildford Press, (2015)
- [8] Fahim, S.M. Siddiqui, K., Anjam, M., Aziz, F., (2013) "Public Opinion towards Advertising: Factor Analytic Findings From Pakistan", *European Journal of Business and Management*, **5**(6) pp:175-180
- [9] Siddiqui, K., "Personality influences on Customer Satisfaction." *African Journal of Business Management*.**6**(11) pp: 4134-4141(2012)
- [10] Siddiqui, K., "Personality influences Mobile Phone usage", *Interdisciplinary Journal of Contemporary Research in Business*,**3**(3) pp:554-565 (2011).
- [11] Siddiqui, K., Personality Influences Consumer Behaviour, *LAP-Publishing*, Germany (2013).
- [12] Wind, Y., Green, P.E., and Jain A.K., "Higher Order Factor Analysis in the Classification of Psychographic Variables," *Journal of the Market Research Society*, October (1973).
- [13] Siddiqui, K., "Personality influences Customer switching", *Interdisciplinary Journal of Contemporary Research in Business*, **2**(10) pp: 363-371, (2011).

- [14] Siddiqui, K., "Personality influences Advertisement Likeness", *Interdisciplinary Journal of Contemporary Research in Business*, 3(2) pp: 631-638 (2011).
- [15] Siddiqui, K., Shahid Bashir, Sarki, I.H., Jaffari, S.I.A., & Abbas, S., "Intergenerational influences on Brand preferences: Consumables VS Durables". *International Journal of Business and Social Science*, 3(14) pp:840-847 (2012)
- [16] Wind, Y., Green, P.E., and Jain A.K., "Higher Order Factor Analysis in the Classification of Psychographic Variables," *Journal of the Market Research Society*, October (1973).
- [17] Siddiqui, K., "Heuristics for Sample Size Determination in Multivariate Statistical Techniques", *World Applied Sciences Journal*, 27(2): pp 285-287, (2013).
- [18] Kline, R., "Principles and Practice of Structural Equation Modeling", 3rd Edition, NY: Guilford Press, (2010).
- [19] Comrey, A.L., & Lee, H.B. A first course in factor analysis, 2nd Edition. Hillsdale, NJ: Lawrence Erlbaum Associates, (1992).
- [20] Thompson, B. Exploratory and confirmatory factor analysis: Understanding concepts and applications. Washington, DC: American Psychological Association (2004).
- [21] MacCallum RC, Widaman KF, Zhang S, Hong S. "Sample size in factor analysis". *Psychological Methods*. 4(1) pp:84-99. (1999)
- [22] Meyers, L.S., Gamst, G., Guarino, A.J., Applied "Multivariate Research: Design and Interpretation", Sage Publications (2006)
- [23] Tabachnick, B.G. & Fidell, L.S. Using Multivariate Statistics. 3rd ed. NY: Harper-Collins (1996)
- [24] Hair, J.F., Black, W.C., Babin, B.J., Anderson, R.E., "Multivariate Data Analysis", 7th Edition, NJ: Prentice Hall, (2010)
- [25] Hair, J.F., Black, W.C., Babin, B.J., Anderson, R.E., "Multivariate Data Analysis", 7th Edition, NJ: Prentice Hall, (2010)
- [26] Thompson, B. Exploratory and confirmatory factor analysis: Understanding concepts and applications. Washington, DC: American Psychological Association (2004).
- [27] Byrne, B. M., "Structural Equation Modelling With AMOS: Basic Concepts, Applications, and Programming" (Multivariate Applications Series), London: Lawrence Erlbaum Associates (1998)
- [28] Hair, J.F., Black, W.C., Babin, B.J., Anderson, R.E., "Multivariate Data Analysis", 7th Edition, NJ: Prentice Hall, (2010)
- [29] Hair, J.F., Black, W.C., Babin, B.J., Anderson, R.E., "Multivariate Data Analysis", 7th Edition, NJ: Prentice Hall, (2010)