

Making the Case for Using Exploratory Structural Equation Modeling Within Biomedical Research

Hyeri Hong¹ and Walter P Vispoel^{2*}

¹California State University, Fresno, USA

²University of Iowa, USA

*Corresponding author: Walter P Vispoel, University of Iowa, USA

ARTICLE INFO

Received: 📅 March 22, 2024

Published: 📅 April 04, 2024

Citation: Hyeri Hong and Walter P Vispoel. Making the Case for Using Exploratory Structural Equation Modeling Within Biomedical Research. Biomed J Sci & Tech Res 55(5)-2024. BJSTR.MS.ID.008778.

ABSTRACT

In this brief article, we summarize the characteristics of three approaches to factor analyzing data: exploratory factor analysis, confirmatory factor analysis, and exploratory structural equation modeling, and describe results from an empirical study that favored the use of exploratory structural equation modeling over the other approaches. We also direct readers to resources with computer code and further details for conducting the illustrated analyses.

Abbreviations: EFA: Exploratory Factor Analysis; CFI: Comparative Fit Index; TLI: Tucker-Lewis Index; RMSEA: Root Mean Square Error of Approximation; ESEM: Exploratory Structural Equation Modeling

An Overview of Factor Analytic Techniques

Since its inception in the early years of the 20th century (Spearman [1,2]), factor analysis has been used extensively in applied research across numerous disciplines. The fundamental purpose of factor analysis is to establish the number and nature of latent variables or factors that explain associations among observed scores. A factor is an unobservable variable that affects more than one observed score and accounts for correlations among those scores. Common applications of factor analysis are to determine whether interrelationships among observed indicators can be accounted for by a smaller number of underlying latent constructs (Brown [3]), investigate convergent and discriminant validity for such indicators and constructs (Millon [4]; Van de Vijver & Leung [5]), serve as a method of data reduction (Cox, et al. [6]), and provide a mechanism for investigating and testing theoretical models (Matsunaga [7]). Over the years, Thurstone's [8] "simple structure" common factor model has attracted the most attention. Within this model, each observed score or indicator is represented by one or more common factors and an error term. The variability of each indicator score is divided into two parts: common

variance shared among indicators (i.e., communality) and unique variance that is specific to the indicator or due to random measurement error (i.e., uniqueness). The common factor model encompasses two primary types of factor analyses: Exploratory Factor Analysis (EFA) and Confirmatory Factor Analysis (CFA; Joreskog [9,10]). The goal of both approaches is to use a reduced number of distinguishable latent variables or factors to explain observed associations among the indicators examined (Brown [3,11]).

Exploratory Factor Analysis (EFA)

EFA is driven by data in the sense that it does not require that the number of factors or the pattern of relationships among latent factors and indicators be explicitly specified in advance. EFA is typically used in preliminary research to determine the number of common factors needed and identify which observed scores are the best indicators of the latent factors in relation to factor loadings that represent the relationships between observed scores and latent factors. The number of common factors can be determined based on a variety of methods (see, e.g., Brown [3,11]). Ideally, each observed score would be more

strongly linked to one factor than all others. However, EFA has shortcomings including its emphasis on empirical data over theoretical modeling and the absence of fit indices to evaluate the adequacy of the factor models in explaining interrelations among observed scores. The goal of CFA is to mitigate these drawbacks.

Confirmatory Factor Analysis (CFA)

CFA is an extensively used structural equation method for testing theoretical models that represent relationships between observed scores and latent factors. In contrast to EFA, CFA is theory-driven and highly restrictive in that the researcher must specify the number of factors and how observed indicators relate to those factors. In most common applications of this method, researchers allow each indicator to load on one targeted latent factor but not on other factors to achieve the clearest simple structure. CFA is suitable for directly testing theoretical models and is more parsimonious than EFA (Brown [3,11]).

Exploratory Structural Equation Modeling (ESEM)

ESEM is a more recent approach to factor analysis (Asparouhov Muthén, et al. [12-15]) intended to overcome the highly restrictive nature of CFAs by allowing indicators to have non-zero (i.e., weak, or typically negligible) loadings on non-targeted factors, while still retaining the advantage of CFA in directly testing prespecified theoretical models. ESEM is more data-driven than CFA by allowing indicator scores to load on all factors but expecting those scores to load noticeably higher on targeted than on non-targeted factors. Recent studies have revealed that, when samples are of adequate size, ESEMs pro-

duce better model fits and more precise parameter estimates than do CFAs (Asparouhov & Muthén [12, 16-22]).

An Empirical Example

To illustrate advantages of ESEMs, we include selected results from a recent dissertation study by the first author (Hong [23]) in Table 1. The data reported represent responses from 447,500 residents in the United States (39% male, 61% female; mean age = 24.93), who completed the International Personality Item Pool NEO 120 questionnaire (IPIP-NEO-120) that we obtained from a publicly accessible university website (<https://osf.io/tbmh5/>) created to enhance research into personality-related constructs (Johnson [24]). The IPIP-NEO-120 has 120 items that measure the Big Five personality domain constructs: Agreeableness, Conscientiousness, Extraversion, Neuroticism, and Openness to Experience. Each domain scale has 24 items with six nested 4-item facet subscales (see the note to Table 1 for names of all facets within each domain). We conducted separate correlated multifactor CFA and ESEM analyses for each personality domain with factors corresponding to the six facets included within a given domain. In Table 1, we report three model fit statistics for each analysis. These include the Comparative Fit Index (CFI), Tucker-Lewis Index (TLI), and Root Mean Square Error of Approximation (RMSEA). In keeping with guidelines suggested by Hu and Bentler [25], we considered fits to be respectively acceptable and excellent with values of 0.90 and 0.95 or higher for CFIs and TLIs and values of 0.08 and 0.06 or lower for RMSEAs. As can be seen in Table 1, ESEMs provided noticeably better model fits than CFAs in all instances and matched or exceeded all criteria for excellent fits with only one exception (the TLI = 0.947 for Extraversion).

Table 1: Fit Statistics for CFAs and ESEMs and Average Off-Target Loadings for the ESEMs.

Domain	CFA Fit Statistics*			ESEM Fit Statistics and Average Off-Target Factor Loadings			
	CFI	TLI	RMSEA	CFI	TLI	RMSEA	Average off-target loading
Agreeableness	0.906	0.891	0.055	0.983	0.968	0.029	0.026
Conscientiousness	0.942	0.933	0.046	0.991	0.983	0.021	0.032
Extraversion	0.897	0.879	0.063	0.972	0.947	0.042	0.042
Neuroticism	0.922	0.909	0.056	0.976	0.954	0.04	0.037
Openness	0.895	0.878	0.049	0.974	0.950	0.031	0.014
Mean	0.912	0.898	0.054	0.979	0.960	0.033	0.030

Note: Subscale facets for the personality domains described in the table include: Trust, Morality, Altruism, Cooperation, Modesty, and Sympathy for Agreeableness; Self-Efficacy, Orderliness, Dutifulness, Achievement-Striving, Self-Discipline, and Cautiousness for Conscientiousness; Friendliness, Gregariousness, Assertiveness, Activity Level, Excitement-Seeking, and Cheerfulness for Extraversion; Anxiety, Anger, Depression, Self-Consciousness, Immoderation, and Vulnerability for Neuroticism; and Imagination, Artistic Interests, Emotionality, Adventurousness, Intellect, and Liberalism for Openness to Experience. CFA: confirmatory factor analysis; ESEM: exploratory structural equation modeling; CFI: comparative fit index; TLI: Tucker-Lewis index; RMSEA: root mean square error of approximation. All analyses were based on correlated multifactor models using maximum likelihood parameter estimation.

*Within the CFAs, all off-target loadings are set equal to zero.

For CFAs, in contrast, CFIs or TLIs never reached levels for excellent fits and failed to achieve acceptable fits across fit indices within three of the five personality domains (Agreeableness, Extraversion, and Openness to Experience). As would be desired, factor loadings for non-targeted factors in the ESEMs, on average, were negligible in size, ranging from 0.014 to 0.042 over the five personality domains.

Conclusion

Our intent in this brief article was to introduce readers to possible benefits of applying ESEMs within scientific research studies by combining the best aspects of EFAs and CFAs. To facilitate applications of ESEMs, routines for analyzing such models are now available in the computer packages Mplus (Muthén & Muthén [26]) and R (Prokofieva et al. [27]). Examples and computer code for analyzing the multifactor ESEMs described here as well as hierarchical and bifactor ESEMs can be found in Hong [23] and (Hong, et al. [28]). For more in-depth information about the nature and relative advantages of ESEMs over other procedures, we direct readers to the comprehensive treatment of such models in (Marsh, et al. [13]).

References

1. Spearman C (1904) General intelligence objectively determined and measured. *American Journal of Psychology* 15(2): 201-292.
2. Spearman C (1927) *The abilities of man*. London: MacMillan.
3. Brown TA (2015) *Confirmatory factor analysis for applied research* (2nd Edn.), New York, NY: Guilford.
4. Millon T (1994) *Manual for the MCMI-III* (3rd Edn.), Minneapolis, MN: National Computer Systems.
5. Van de Vijver F, Leung K (1997) Methods and data analysis of comparative research. In: J W Berry, YH Poortinga, J Pandey (Eds.), *Handbook of cross-cultural psychology: Theory and method* Allyn & Bacon, pp. 257-300.
6. Cox, Walker JR, Enns MW, Karpinski DC (2002) Self-criticism in generalized social phobia and response to cognitive-behavioral treatment. *Behavior Therapy* 33(4): 479-491.
7. 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.
8. Thurstone LL (1947) *Multiple-factor analysis*. Chicago: University of Chicago Press.
9. Jöreskog (1969) A general approach to confirmatory maximum likelihood factor analysis. *Psychometrika* 34(2): 183-202.
10. Jöreskog (1971) Statistical analysis of sets of congeneric tests. *Psychometrika* 36(2): 109-133.
11. Kline (2016) *Principles and practice of structural equation modeling*. Guilford Press.
12. Asparouhov T, Muthén B (2009) Exploratory structural equation modeling. *Structural Equation Modeling: A Multidisciplinary Journal* 16(3): 397-448.
13. Marsh HW, Morin AJ, Parker PD, Kaur G (2014) Exploratory structural equation modeling: An integration of the best features of exploratory and confirmatory factor analysis. *Annual review of clinical psychology* 10: 85-110.
14. Morin AJS, Marsh HW, Nagengast B (2015) Exploratory structural equation modeling. In: GR Hancock & R. O. Mueller (Eds.), *Structural equation modeling: A second course* (2nd Edn.), Charlotte, NC: Information Age, pp. 395-436.
15. Muthén B, Asparouhov T (2012) Bayesian structural equation modeling: A more flexible representation of substantive theory. *Psychological Methods* 17(3): 313-335.
16. Booth T, Hughes DJ (2014) Exploratory structural equation modeling of personality data. *Assessment* 21(3): 260-271.
17. Mai Y, Zhang Z, Wen Z (2018) Comparing exploratory structural equation modeling and existing approaches for multiple regression with latent variables. *Structural Equation Modeling: A Multidisciplinary Journal* 25(5): 737-749.
18. Marsh HW, Muthén B, Asparouhov T, Lüdtke O, Robitzsch A, et al. (2009) Exploratory structural equation modeling, integrating CFA and EFA: Application to students' evaluations of university teaching. *Structural equation modeling: A multidisciplinary journal* 16(3): 439-476.
19. Marsh HW, Lüdtke O, Muthén B, Asparouhov T, Morin AJS, et al. (2010) A new look at the big five factor structure through exploratory structural equation modeling. *Psychological Assessment* 22: 471-491.
20. Morin AJS, Arens A, Marsh H (2016) A bifactor exploratory structural equation modeling framework for the identification of distinct sources of construct-relevant psychometric multidimensionality. *Structural Equation Modeling: A Multidisciplinary Journal* 23: 116-139.
21. Perera HN (2016) Construct validity of the Social Provisions Scale: A bifactor exploratory structural equation modeling approach. *Assessment* 23(6): 720-733.
22. Reis D (2017) Further insights into the German version of the Multidimensional Assessment of Interoceptive Awareness (MAIA): Exploratory and Bayesian structural equation modeling approaches. *European Journal of Psychological Assessment* 35: 317-325.
23. Hong H (2022) *Finding Optimal Factor Analytic Representations of the Structure of Personality*. Doctoral dissertation, University of Iowa.
24. Johnson JA (2014) Measuring thirty facets of the Five Factor Model with a 120-item public domain inventory: Development of the IPIP-NEO-120. *Journal of Research in Personality* 51: 78-89.
25. Hu L, Bentler PM (1999) Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal* 6(1): 1-55.
26. Muthén LK, Muthén BO (2017) *Mplus: Statistical Analysis with Latent Variables: User's Guide* (Version 8th); Muthén & Muthén: Los Angeles, CA USA 2017.
27. Prokofieva M, Zarate D, Parker A, Palikara O, Stavropoulos V (2023) Exploratory structural equation modeling: A streamlined step by step approach using the R Project software. *BMC Psychiatry* 23(1): Article 546.
28. Hong H, Vispoel WP, Martinez AJ (2024) Applying SEM, exploratory SEM, and Bayesian SEM to personality assessments. *Psych* 6(1): 111-134.

ISSN: 2574-1241

DOI: 10.26717/BJSTR.2024.55.008778

Walter P Vispoel. Biomed J Sci & Tech Res



This work is licensed under Creative Commons Attribution 4.0 License

Submission Link: <https://biomedres.us/submit-manuscript.php>



Assets of Publishing with us

- Global archiving of articles
- Immediate, unrestricted online access
- Rigorous Peer Review Process
- Authors Retain Copyrights
- Unique DOI for all articles

<https://biomedres.us/>